理性之谜

THE ENIGMA OF REASON

［法］雨果·梅西耶（Hugo Mercier）［法］丹·斯珀伯（Dan Sperber）著

张慧玉 刘雨婷 徐开 译



中信出版集团

版权信息

书名:理性之谜

作者:[法]雨果·梅西耶,[法]丹·斯珀伯

译者:张慧玉 刘雨婷 徐开

ISBN:9787508692289

中信出版集团制作发行

推荐序

理性不在风平浪静中，而在众声喧哗时

如果我请你闭上眼睛，为“理性思考”这几个字想象出一幅画面，那么在你脑海中浮现出来的会是什么？

在读《理性之谜》之前，我脑海里可能会浮现两个形象：叼着烟斗、躺在安乐椅上推理罪犯身份的侦探，在黑板上演算公式、论证学术新知的科学家。

没准儿你想到的也是他们，在我们熟悉的各种形象里，恐怕没有比侦探或科学家更擅长调用理性思考来推理真相、解决难题的了。而在我的脑海里，这两个形象还有一个共同点：他们都是孤胆英雄。

福尔摩斯追查线索时也许得拉上华生，但最后把线索拼凑起来，靠的还是他自己头脑里的“思维宫殿”。没有哪个侦探会在进行关键推理时还要拉上助手来帮助自己整理思路。

而你也一定能回想起几个关于科学家独自深思解决科学难题的故事。比如物理学史就有一个关于“孤独天才”的著名故事：德国物理学家海森堡跑到一个海岛上闭关，在独自沉思数月之后以一人之力推导出量子力学的基石——不确定性原理。

理性思考似乎只是“一个人”的事，越是在不受打扰的情况下独自一人深思熟虑，理性的威力就发挥得越充分，对吧？

如果你也这么认为，那你一定得读一读《理性之谜》。读完这本书我才意识到，这种传统的印象可能完全是错的！

事实上，理性思考这种能力对孤胆英雄一点儿也不友善，它最能施展拳脚的场合，反而是在人与人互动之时。

《理性之谜》的两位作者（他们都是顶尖的认知科学家）提出了一个相当颠覆直觉的新理论——理性的互动论来概括这一观点。在互动论看来，人类在进化中获得理性，并不是理性可以帮助人们更清晰地思考问题，而是理性可以帮助人们更有效地互动。

也就是说，理性原本是用来搞社交的。

理性思考解决了社交互动中的两大难题：向别人证明自己是一个靠谱的合作者，以及揭穿那些不靠谱的合作者。为自己的行动辩护，质疑别人的行动不靠谱，这才是理性的两大原始功能。理性的确帮助人类解决了无数难题，而那些成果几乎都是人们在一场场"忽悠"与"反忽悠"攻防战中收获的战利品。

这个理论框架能有效地解释理性的某些奇特属性。比如说，为什么理性有一种奇特的双面性——"宽"以律己，"严"以待人。

当用理性来反思自己的观点和行动时，它似乎总是急于为自己辩护，哪怕其中错漏百出，我们也不容易察觉，轻易就被自己浅薄的推理说服。

侦探们独自在"思维宫殿"中做出的推理可能并不会高明到哪里去。《理性之谜》里有一个关于福尔摩斯的小段子是这样的：

> 福尔摩斯：亲爱的华生，你有那么迷人的女伴在场，我没敢打扰！我已经观察你俩好久了，我强烈建议你再约一下这位女士！她跟你非常般配。虽然她比你年龄小，但也差得不多。我知道，你喜欢深褐色头发的女人，而她是金发，但无疑应该忽视这一差异。我还观察到，你们在谈及私事时，没有任何亲密的肢体接触，但我相

信，你们很快就会有的。华生，我敢保证，这个女人最适合你！

华生：她是我妹妹。

如果福尔摩斯是个真实人物，那么他的推理会多半更像这个段子，而不会像小说中那般神准。当人们独自推理时，更有可能的情况是像这个段子里的福尔摩斯，明明自己的推理中有一些显眼的线索，从中不难推断出华生的女伴并非他的情人，但也许他是太期待华生找到幸福了，心中先认定华生有了女朋友，接着便对所有矛盾的线索视而不见。

为什么会这样呢？在互动论的视角下，原因很简单——当人们独处时，更接近于互动中为自己的行动辩护的情境。这时，人们倾向于宽松的标准，因为此时人们只求合作者接受自己给出的理由，而不求推理是否真的合理。

科学家其实也常犯这种错误。《理性之谜》中引用的一项研究发现，当科学家独自思考自己的研究结果时，他们其实很容易找到各种理由去抹掉那些对自己提出的研究假设不利的数据。当实验结果让人失望时，科学家并不会不偏不倚地去推理、去质疑他们最初的假设可能不对，相反，他们很容易对可以挽救原先假设的各种糟糕理由感到满意，如“结果不好是因为出现了技术问题”“是因为实验的操作有缺陷”等。

这听上去很糟糕吧？可这又怎么解释人类通过科学研究得到那么多靠谱的发现呢？

原因在于，一旦人们把理性思考的对象转向他人，就从“忽悠模式”切换到了“反忽悠模式”，由于轻信不靠谱的合作者可能会让自己付出惨重代价，这时理性思考的标准就一下子严格起来了。

科学界就是利用这一点来强力纠错的。学术圈现在有一个通行的规则——同行评议程序：一篇论文被提交到期刊编辑部之后，都要被转发

到研究者的同行那里接受审查。这道程序说白了就是让同行给你“抬杠”，然后你再“反抬杠”，只有你“反抬杠”成功，论文才能被发表。

每当跟别人“抬杠”的时候，人们的理性思考就忽然变得敏锐起来，那些靠自己在独自思考时的理性难以发现的错误，往往在这时会被他人“抬杠”时的理性发现。其实，在真实的科学史上，孤独天才式的科学家故事实属罕见。绝大多数重要的科学发现，都是在高质量的争辩中获得的。

关于海森堡发现不确定性原理的故事，如果我们仔细去审查史料，就会发现他解决问题的整个过程其实充满了互动。其中最重要的是他一直与另一位大物理学家泡利保持着非常紧密的通信联系。正是因为不停地面对泡利提出的疑问，海森堡才能把他的想法充分发挥出来。事实上，海森堡不确定性原理的论文初稿最早就出现在给泡利的一封信里。即使是天才，也需要与其他人争辩，才能发展出他们最好的想法。

在前面那个段子里，如果福尔摩斯的死对头莫里亚蒂教授也在场，而那一番蹩脚的推理是从莫里亚蒂教授口中说出的，那我想福尔摩斯绝对能从中挑出一堆漏洞。

这就是为什么理性思考并不适合在孤胆英雄一个人的头脑里发生，当头脑与头脑交锋时，理性才能发挥出全部威力。

理性的这种特点，让我想起这些年在创业圈中很火的一个概念——精益创业。

精益创业是这样一种理念：企业研发产品时切忌自己关起门来，把产品打磨得尽善尽美再推出市场，这很像福尔摩斯独自一人绞尽脑汁推理半天，结果是吃力不讨好。企业正确的做法是用最快的速度研发出一款基本可用的原型产品，并迅速投入市场，让市场的反馈告诉它们产品哪些地方最值得改进，最后在反馈的基础上快速迭代产品。这些年来很

多创新企业成功地利用这种思路迭代产品。比如微信，刚开始推出来的是一个功能极其简陋的版本，品质远不如竞品，但在几轮迭代之后就进化到远远甩开对手的水准。

福尔摩斯：亲爱的华生，我好像没有见过你这位迷人的女伴，可否向我介绍一下……

华生：她是我妹妹。

精益创业其实就像这样一来一去的简单互动，在互动中充分利用各方优势，高效省力地解决问题。

大道归一，理性互动论的视角其实还可以迁移到生活中很多案例之中，你可以在《理性之谜》中细细体会。

读完《理性之谜》，我再想起“理性思考”，脑海中浮现出的还是两个形象：福尔摩斯叼着烟斗……与莫里亚蒂斗嘴，科学家在黑板上演算公式……以便拿出证据反击同行的“抬杠”。

人的理性，不在风平浪静中，而在众声喧哗时。

魏知超

宁波大学心理学系教师

喜马拉雅FM《魏知超：心理学新知课》主理人

序言
双重谜团

它们会吃喝拉撒，会睡觉打鼾，会出汗，会发抖，有性欲时会交配，它们的生与死都混乱不堪。这就是动物，人类也是动物！然而上天赐予了人类理性。理性使人类与众不同，并凌驾于其他生物之上——一般来说，西方哲学家如是宣称。

人类视兽性为耻辱和丑闻，这至少应包含在人类诉诸理性的过程之中，这种诉诸理性的能力使人类变得博学而明智。这种机能是理性而非语言，因为其他动物似乎也有某种形式的语言；这种机能更非灵魂，因为“灵魂”这种说法太神秘了。上天赐予人类理性之后，人类虽仍为动物，但已非野兽。

理性：有缺陷的超能力？

达尔文（Darwin）发现，人类这一物种共有的任何特性都并非上天的恩赐，而是生物进化的结果。理性这种特性必然经历了进化过程。要不然呢？自然选择不是孕育出了许多惊人的机制吗？

以视觉为例。大部分动物物种都获益于视觉这一惊人的生物适应特性。视觉把专用的外部器官（眼睛）与大脑的特定部分相连接，让动物可以从视网膜刺激模式中提取出有关远处物体的特性、位置及运动状况的精准信息。这项任务极其复杂，无论怎样都比理性复杂得多。人工智能领域的研究员在视觉及推理的建模、施行上付出了艰苦卓绝的努力。尽管如此，机器视觉仍处于起步阶段，远不及人类视觉。另外，许多计

算机推理模型都宣称自身性能要比人类理性更为优越（过于乐观了）。因此，若视觉可以进化，那么理性也可以。

有人认为，优于视觉的理性是一项多用途的机能。它将人类认知提升到了一个新的高度。没有理性，动物认知就会受到本能的束缚，知识与行为都会受到极大的限制。认知一旦得到理性的加持，就可以让人加深对各个领域的了解，并根据新奇又富有雄心的目标来调整自身行动，这大概就是一般的说法。但是，等一等，如果理性是这样一种超能力，那它为何与视觉不同，有且仅有人类才得以进化这一特性呢？

的确，某些显著的适应特性实属罕见。只有少数物种，如蝙蝠，拥有发展完备的回声定位系统。蝙蝠发出超声波后，其周围环境会给出回声，蝙蝠便是利用回声来识别并定位障碍物或移动猎物的。其他大多数动物都不会做类似的事情。

视觉与回声定位能力有诸多共同特性。狭隘的辐射范围（视觉中的光线、回声定位能力中的超声波都是如此）能提供诸多信息，这些信息与丰富多样的认知、实践目标相关。可为何视觉如此常见，而回声定位能力如此罕见呢？这是因为，在大多数情况下，视觉更有效。只有在不可见或视觉严重受损的生态位中，回声定位能力才具有适应性——像蝙蝠那样住在洞穴中并在夜间猎食就是很好的例子。

理性如此罕见（可以说为单一物种独有），是因为其适应性只体现在人类所在的这一十分特殊的生态位中吗？这一有趣的可能性值得探讨。然而，这与标准的推理方式格格不入。标准的推理方式认为，不论理性的运作环境如何、目的是什么，理性终究会提高人类的认知水平。要了解为何只有少数物种具备回声定位的能力很简单，但要了解为何只有人类拥有理性则更有挑战性。

以轮子为例。动物并没有轮子，为何没有呢？[注]毕竟建造装有轮子的车辆要比建造有腿或有翅膀的动物容易多了（正如形成推理模型似乎

比形成视觉模型要容易得多）。然而，人工轮子是分开制造的，之后才装到车辆上去，而生物轮子则必须在其原位生长。既然身体上这个能自由转动的部分通过神经及血管与身体其他部分相连，它又怎能在联系如此紧密的情况下单独发挥作用呢？切实可行的生物学解决方案构思起来没那么容易，这还只是一部分问题。

一个复杂的生物适应特性要进化，必须经历一系列进化步骤，从初步的预兆到发展完备的机制，这一过程中的每一次改变都是自然选择所青睐的（至少没有被淘汰）。举例来说，昆虫、软体动物或哺乳动物的复杂视觉系统，起初只是感光细胞，而后经历了漫长的改变之旅，其中每一次改变都是具有适应性的，至少是温和而平静的，最终才进化成今天的模样。若是假定动物有可能从没有轮子进化到有轮子，那么其间必然经历了一系列类似的适应性步骤，但这些步骤发生的可能性简直微乎其微，所以可断定其从未发生过。

这样看来，给动物认知附以理性的加持或许就像给动物运动附以轮子的帮助一样，都是极其不可能的进化结果。理性如此罕见，或许是因为其必须历经一系列可能性很小的步骤才能进化，且因为其只在进化史上发生过一次，离现在还很近，所以目前只有我们这一个幸运的物种从中获益。

理性逐步进化的一系列步骤仍然是个谜。与把超人或蜘蛛侠的超能力融入原本普通的人类特性相比，把理性融入人类更普通的认知能力似乎更难。当然，可能有人会说，理性是一种移植物、一组附加组件、一套文化装置（古希腊就有人认为理性是一种发明创造），而不是一种生物适应特性。但本身并不具备理性这种超能力的物种，怎么可能去发明理性呢？虽然理性显然受益于各种文化成就，但人类这一物种制造、评估并运用理性的能力迫切需要进化角度的解释。唉，可能解释后的结果也不过是失望而已。

让这一问题更糟糕的是：挥手本身似乎指向了一个错误的方向。对

照之举便是设想生物轮子冲破万难在一个动物物种中得到进化。我们并不知道这种进化是如何发生的。不过，若这些轮子能让动物在自然环境中高效运动，我们就能知道它们为何会进化至今了，换句话说，就能够知道其功能了。我们或许会想，动物轮子与所有的生物器官一样会有弱点，其功能会偶尔失灵。但我们不会认为这种运动系统存在系统性缺陷，进而让其功能发挥得不尽如人意——举例来说，若两侧的轮子大小不一，动物就很难朝着既定方向前进。有的生物机制被认为是“不适应的适应特性”，很可能只是因为表述有误。理性的标准表述便是如此。

心理学家声称已经证明人类理性是有缺陷的。有人认为理性的工作做得很糟糕，这一观点已成了老生常谈。心理学家及哲学家通过反复的实验确定，人类会在推理中犯极其严重的错误。而这并不仅因为人类不善于推理，还因为他们存有系统性偏见。理性之轮失去了平衡。

除了这一陈词滥调，论战也已爆发。既然理性存在缺陷，那么缺陷究竟有多严重呢？推理中的成功与失败该如何衡量呢？负责的机制是什么呢？尽管论战各方的分歧通常很尖锐，但也没人去质疑基本的教条。所有人都理所当然地认为推理的任务就是帮助个体获取更多的知识，做出更好的决策。

你如果接受这个教条，就会感到迷惑不解——理性竟会有失公允、不合逻辑。自相矛盾的是，推理竟然常常让人难以统一意见，反而让人加深分歧。那起初为何要接受这个教条呢？好吧……可能是因为惯例的包袱太重了。你或许会问推理还有什么功能。

人们一般认为理性具有双重谜团。理性并不是一项普通的心理机制，而是进化（过去是上帝）赋予人类的认知超能力。仿佛这一点还不够神秘，超能力竟然是有缺陷的。理性渐渐把人引入歧途。理性是有缺陷的超能力吗？这是真的吗？

我们的目标就是解开这个双重谜团。我们会证明理性是如何与个人

心智、社会互动及人类进化相适应的。因此，我们向传统观念发起了挑战，拒绝接受教条，重新思考理性的机制与功能。

我们将去往何方

2 000多年来，哲学家一直致力于研究理性，其间，有50多年用于对推理开展集中性实验研究。自古以来不乏最伟大的思想家为这项研究披荆斩棘。如果不是因为哲学与心理学传统在内部一直存在激烈的竞争，那么判定这类想法大多数都走上了错误的轨道，是毫无根据的。

在指引人们获取真理知识、做出明智决定的路途上，理性发挥了多大的作用呢？人类又有多善于利用理性呢？对于这些古老的辩论，我们不想多加赘述以至于令人费解，大概最近在心理学家的参与之下，这些古老的辩论已升级到“理性战争”的地步了。但是，我们会在本书第一部分“撼动教条”中，挑选出一些矛盾和冲突，这些矛盾和冲突能够揭露的就是标准推理理论存在的问题有多严重，相应的解决方案又有多不足。我们会指出，参与这些激烈辩论的各方已经成功地削弱了彼此的力量，以至于最佳方案大概就变成，从辩论场地收集任何可能仍有作用的东西，然后携其去更广阔的天地间开始新的冒险。

总之，比起证明摇摇欲坠的想法是错误的，我们更感兴趣的是，对理性形成新的、可以解开双重谜团的科学性认识。本书将证明，理性远非奇怪的认知附加组件，也非某些不太可能的进化恰巧赋予人类的超能力，而十分自然地与人类其他认知能力相适应，尽管有明显证据表明两者之间的适应并不自然，理性还是很好地适应了自己真正的功能。

要了解理性如何进化、如何运作，要注意的不仅是什么使理性如此特别，还要注意理性是如何与其他心理能力相互适应的，以及与之有多少共同点。进行推论时会用到许多机制，理性只是其中一个。在第二部

分“理解推论”中，我们把理性与其他推论机制联系起来考虑，见图0.1。

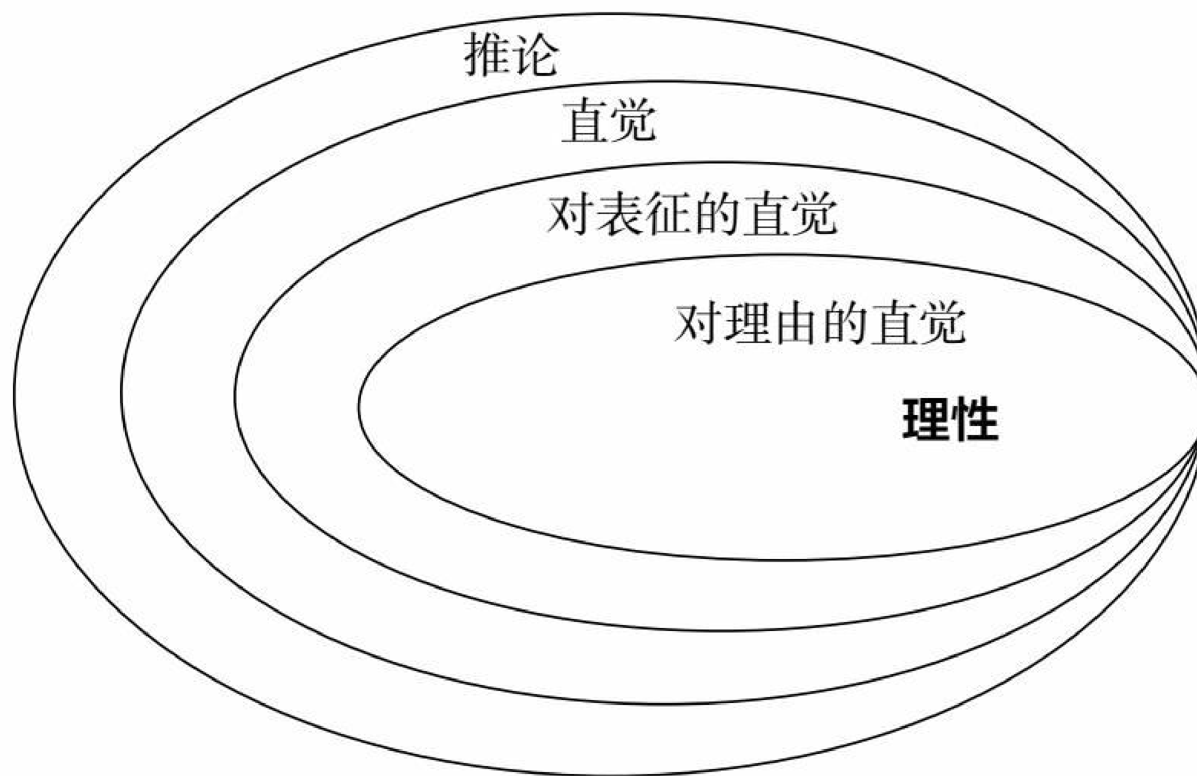


图0.1 理性是如何被嵌入好几个推论类别下的

动物一直都在进行推论：它们利用已知事物来推断未知事物，如预测接下来会发生什么并采取相应的行动。动物也把此举应用到了某些一般性推论能力中吗？当然没有。但是，动物用到了许多不同的推论机制，每种机制都处理截然不同的问题，如吃什么，与谁交配，何时攻击，何时逃跑，等等。

人类与其他动物相似，不仅拥有一种通用性推论能力，还能运用其他各种专门性的推理机制。然而，对人类而言，其中许多机制并非本能，而是孩童在成长过程中通过与他人互动习得的。尽管如此，这类习得性机制大多数都有本能性的基础。举例来说，不论是说沃洛夫语（Wolof）、英语，还是说塔加洛语（Tagalog），人们都并非天生就会，然而，对说话声音加以特别的留意、习得并遵循某一群落的语言需要的就是本能性的基础。

众所周知，其他动物都是在无意识的状态下表现出自身的推论能力的。人类同样也会无意识地表现出多种推论能力，学习母语就是一个例子。不过，也有许多推论是人类在某种程度上的有意识之举。接下来，我们会讨论直觉。当你有了某种直觉——举例来说，你凭直觉知道好友莫莉（Molly）现在很沮丧，尽管他并没有这么说，甚至可能否认这一点，但这个直觉还是突然出现并完全成形。不过，与此同时，你认为这种直觉是身体内部自发形成的，是思维内部得出的结论。直觉就如同思想铸成的巨大冰山：或许我们能看到的就只有冰山之尖，但我们知道，冰尖之下还有无尽神秘亟待探索。

许多观点都是关于思维的，如丹尼尔·卡尼曼（Daniel Kahne-man）的著作《思考，快与慢》（*Thinking, Fast and Slow*）㊟，这类思索围绕直觉与推理之间的差异展开，仿佛二者是迥然不同的推论形式。我们的观点恰恰相反，我们坚持认为，推理本身就是一种直觉推论。

其实，在一般性直觉与特殊性推理之间存在中间范畴。人类不仅能表现环境中的事物及事件，还能对自身表征进行表现。我们不仅能产生对他人所思所想的直觉，还能产生对抽象观念的直觉。这些由表征而产生的直觉对我们互相理解沟通、分享观点与价值观有重要的作用。本书会证明，理性是对一种表征（即理由）进行直觉性推论的机制。

在第三部分“反思理性”中，在某些重要的方面，我们选择了与主流相悖的方式，即摒弃了将理性与直觉进行对比的标准化方式。我们把理性（reason，作为心理机能）研究与理由（reasons，作为正当理由）研究归为相同的话题。然而，在哲学及心理学上，二者被视为截然不同的领域。

尽管人们一般把理性当作可以让个体思考得更周全、更高效的方式，然而我们却认为个体主要是在与他人的互动中运用理性。我们拿出理由是为了在他人面前为自身观点及行动辩护，同时也是为了作为说服

他人的论据，来让他人按自己的想法进行思考、做出行动。我们运用理性也不是为了对自身看法进行评估，而是为了对他人所给的理由进行评估，他人会运用这些理由来为自己的观念辩护，进而说服我们。

尽管人们一般认为理性是逻辑应用，或者至少是某种规则系统，可以用来拓展知识、改善决策，然而我们却认为理性更擅于投机取巧、博采众长，还不会受到正规范式的束缚。我们认为，逻辑在推理中的主要作用很可能是修辞，因为逻辑能够对直觉性论据进行简化、系统化，从而突出强调直觉性论据的说服力。

那么，理性为何会进化？除了其他更普通的推论形式所提供的，理性还提供了哪些仅对人类具有特殊价值的东西呢？为了做出回答，让我们从更广泛的视角来看这个问题。

我们认为，理性具有两项主要功能：一是拿出为自身辩护的理由，二是拿出说服他人的论据。这两项功能密切相关，用到的理由或论据属于同一类型。

为何要费力解释并为自身辩护呢？人与其他动物的不同之处不仅表现在人拥有发展得更为良好的认知能力，关键还在于人知道去合作，会把握尺度。人类不仅跟亲属合作，也跟陌生人合作；不仅在迫在眉睫的时候合作，还在追求长远目标时合作；不仅有物种特有的联合行动形式，还会共同建立新的合作形式。人类这样的合作会催生出其特有的协调及信任问题。

理性的第一项功能就是促成人类丰富多样的合作形式。人类会举出理由来解释自己、证明自己，达到为自己辩护的目的，也会指出自己眼中能够佐证自身观点、证明自身行为的事物。他们这样做就能够让他人知道自己所能做到的程度，并婉转表达自己希望从他人那里得到的东西。在决定相信谁并实现协作的过程中，对他人的理由进行评估显得意义重大。

人类与其他动物的区别还在于，人类与他人分享的信息更为丰富、更为广阔，也在更大程度上依赖于这种沟通和交流。为了成为富有竞争力的成年人，我们每个人都需要向别人学习很多东西。我们的技能及常识与其说是个人经验，不如说是社会经验。在大多数日常活动中（家庭生活、工作、爱情或休闲娱乐），我们十分依赖学来的东西。我们从沟通交流中获取这些不可或缺的巨大利益，与此同时，我们也失去了对大量错误信息的抵抗能力。在听取他人意见时，我们想要的是诚实无欺的信息。在与他人交谈时，我们却又经常根据自身兴趣来误导他人，虽不一定直接撒谎，但至少会歪曲、删减或夸大事实，从而更有力地影响他人的看法及行为。

因此，在听取他人意见时，我们应该明智地信任对方，有时也有所怀疑。在与他人交谈时，我们经常需要克服对方情有可原的信任缺失。当对方不值得信任的时候我们才表示怀疑，这可能是最理想的状况了。不过，我们通常会出于谨慎而保留信任，并非因为知道对方不值得信任，而是因为不确定是否应该信任对方。这种沉默或许是明智的——谨言慎行不吃亏，轻率莽撞必后悔——尽管如此，我们还是错过了有价值的信息。无论是对说话者而言，还是对倾听者而言，沟通都是有益的，但我们常因缺乏自信而闪烁其词。

我们认为，即便沟通者在其听众眼中缺乏足够的可信度，不能被不加深究地信任，理性的第二项功能（通过推理与论证加以实现的功能）也能确保有效沟通。沟通者运用理性拿出作为论据的理由来说服谨慎的受众。同样，谨慎的受众也能运用理性来评估这些理由，并接受合理的论据，否决不合理的论据。

我们先前主要研究理性的这项论证功能，并形成了“论证推理理论”。[注]我们在本书中拓宽了视野，既分析了理性的论证功能，又分析了理性的辩护功能，并形成了应对这些机制及理性的两项功能的互动方式。

在第四部分“理性有所为、有所不为”中，我们对理性的作为进行了回顾。在这一部分，我们展示了互动论视角是如何帮助解释理性为何有此作为的，并再次讨论了理性的某些已被人接受却明显说不通的弊端，例如证实偏见（the cofirmation bias）。同时，我们对理性的某些被人忽略的优点进行关注。

进行回顾是因为看到了人类理性存在偏见而又很懒惰。“存在偏见”是因为理性一边倒地找寻支撑推理者观点的正当理由及论据，“懒惰”则是因为理性几乎不去评测自己找到的“正当”理由及论据到底是否“正当”。举例来说，设想有个人喜欢去沙滩度假，在选择下次度假地点的时候，此人自然而然会选阳光明媚的傍海之地，并为此搜寻理由，甚至将那些明显站不住脚的理由也收入囊中（例如，去那个地方的航班正在打折，虽然实际上去往其他许多地方的航班也有同样的折扣）。

人们单独运用理性会产生两种典型结果。若推理者一开始就打定了主意，那么他想到的所有理由都是为了支撑这一决定。这样的话，他就不太可能会改变主意，甚至可能变得过分自信，进而更加偏执。但有时，推理者一开始并不会打定主意，或持有互相抵触的观点。在这种情况下，理性会驱使推理者做出碰巧更易自证的决定，虽然这有时并非最佳选择。设想推理者要在拜访“可怕”的姻亲（尤指公婆或岳父母）与去海滩度假的先后顺序上做决定，先度假后拜访姻亲辩护起来更容易些。理性会驱使推理者做出看起来明智的决定，即选择那个更容易找到理由来自证的选项。然而，如果他一开始就去看望姻亲，而不是想到若要去拜访姻亲就没了去海滩度假的兴致，那他度假归来时可能心情会更好一点——总的来说是更好的选择，却需要付出难以自证的额外代价。

大多数心理学家都意识到，人们的理性有偏见且懒惰，所以常常无法纠正错误的直觉，有时还会让事情变得更糟。但大多数心理学家坚持认为，理性的主要功能是提高个人认知——一项它执行得非常糟糕的任务。另外，互动论视角第一次为理性常遭谴责的偏见与缺点提供了进化

角度的合理解释。

我们将证明，对于一个旨在为自身辩护并说服他人的认知机制来说，存有偏见且懒惰是说得通的。孤独的推理者推理不成功，是因为他在“非正常”的情况下运用理性。你不会指望买来的钢笔在水里也能正常写字，也不会觉得自己胸腔里的肺在溺水后还能正常呼吸。同样，将理性从其进化的互动语境中抽离出来，则无一物可以确保它会产生适应特性。

那么，若理性被放回“正常”的运行环境，若理性要在人们讨论的过程中，即在人们具有交流充分的论据和正常的理由之时才开始发挥作用，那么会发生什么呢？在这种情况下，理性充分施展了其进化而来的功能。值得一提的是，若人们虽存有分歧，但都想找寻真相或问题解决方案，愿意去交流论据，那么最好的想法往往会脱颖而出；不论起初持有该想法或在讨论中萌生此想法的人是谁，他都有可能说服他人。这一结论听起来可能过于乐观，但有大量证据可以支撑，如讨论逻辑问题的学生、协商的陪审团、努力预测下次战争爆发地点的预报员。

在最后三章（第五部分“狂野的理性”）中，我们会展示先前所述的理性的特性及效果有多强大。我们发现孤独的推理有偏见且懒惰，而论证却十分高效。这不仅体现在过分好辩的西方社会，还体现在所有文化类型中；不仅体现在受过教育的成人中，还体现在幼童中。当在道德及政治问题上应用理性时，在听到“理性通常有偏见且懒惰”时，几乎没有人会感到吃惊。相反，在道德及政治领域产生的惊奇，或许可以显示论证工作的高效率，让参与者能够做出更精准的道德判断，让市民能够想出更开明的意见。然而，这些发现只会出现在互动论的视角之下。

最后一章（“孤独的天才”）是关于科学的。通常，我们认为科学是人类理性的巅峰。科学在许多方面都非常杰出，但科学家的理性也如此杰出吗？从牛顿（Newton）、达尔文到爱因斯坦（Einstein），科学通常都靠孤独的天才来取得进步。众所周知，这些天才的理性更为杰出，

那些困扰着普通大众的缺陷在天才那里并不存在。这些天才不仅不需要与他人论证就能想出新理论，论证反而还有可能阻碍他们前进的道路，误解他们革命性的领悟，被并不那么合格的“先锋”嗤之以鼻。最好还是等不那么偏激的下一代人来见证曙光的到来。科学家设法应付所有人类都在运用的、同样有偏见与局限的理性，但科学家也同样获益于理性的优点，尤其是在运用理性更加高效地评估论据之时。

那么，在这十八个章目中，我们推荐的是互动推理理论，这一理论与标准的唯理智理论形成对照。我们坚持认为，理性首先是社会性能力。我们并不否认理性能够带来巨大的智力效益，科学的例子就很好地证明了这一点；相反，我们解释了理性是如何做到这一点的：通过与他人的互动。

你恐怕不会因为我们这么说就信了，因此会有一些论据来供你自行评估其价值。你将看到，当理性成为对理由进行直觉性推论的机制时，它是如何解决双重谜团的前半部分的：理性并不是不可思议地被移植到动物头脑中的超能力，而是在发育特别出众的心智中完美融合的组成部分——理性使人类这种动物有了独特的一面。

为了解决双重谜团的后半部分，本书将会论证，偏见作为理性糟糕的弊端之一，是如何很好地适应了理性论证功能这一特征的。互动论视角之下产生了许多有关人类理性的预测，这些预测有时会令人大吃一惊，而我们给出的证据能够对上述预测进行证实。我们希望通过论据的力量来让你相信互动论视角的正确性，或者至少是在正确的道路上。当然，这也使本书成为其所持观点的一个有力例证。

---

1. 更多详细信息，可参见：Dawkins 1996。
2. Kahneman 2011.
3. Mercier 2016a; Mercier & Sperber 2011.

第一部分
撼动教条

理性是一项给予人类更高等知识与智慧的能力吗？经过50年来对理性的实验性研究，这个在西方传统中占据主导地位的观点已从根本上动摇。在第一到二章中，我们将会展示，旧有教条是如何被撼动的，但又撼动得不够。如今有关推理的主流观点［“双重过程”（dual process）或“快与慢的思考”（fast and slow thinking）］无论多么吸引人，终究只是旧有思想崩塌情况下的权宜性构建。

第一章
理性受审

1619年深秋，巴伐利亚公爵（Duke of Bavaria）麾下军队中的一名志愿者，年仅23岁的勒内·笛卡儿（René Descartes），发现自己在德国南部有大把时间可以挥霍，而身边却没有一个可以说话的人。正如笛卡儿在自己的著作《方法论》（*Discourse on Method*）[注]中回顾的那样，他在一个如炉火般炙热的房间里，终于将自己惊人宏伟的计划构筑完毕。在这个计划中，他要摆脱所有的意见，摆脱所有从别人身上学来的思想，并且从头开始，一步一步重新建造自己的知识体系。理性将成为他唯一的向导。只有当自己确信无疑时，他才能够接受新东西。

笛卡儿对集体性成就表达了一种不屑与鄙视，以此来佐证自己抛弃从他人身上学到的一切知识是合理的。他坚持认为，最优秀的工作应当是由一位大师独立完成的。笛卡儿认为，从书本上学到的东西，“不够接近真理，就像是由许多不同的人的意见组合而成的，像任何具有良好感官的人，都能以自己既有的眼界来推出有关事物的最简单的论断”。[注]

笛卡儿一定会对现今流行的“群体智慧”观念嗤之以鼻。他承认的唯一智慧（至少在科学上如此）是个体理性的智慧：“为了能够从一事物推出另一事物，一个人只要能够不再相信并不真实的东西，坚守正确的规则，那么在还未到达终点时一切就还十分遥远，在还未发现真理时一切就还深深隐藏其后。”[注]

笛卡儿为何决定只相信自己的智力呢？他相信自己具备独一无二的

推理能力吗？相反，他坚持认为“每个人都自然地、平等地拥有准确的判断力与鉴别真假（事实上被称为良好的感觉或理性）的能力”。[注]但如果人类真的拥有这种鉴别真假的能力，那为何我们会在到底什么才是真的这一问题上分歧如此之大？

“最有才智的人能做最好的事，也能做最坏的事”

我们大多数人都觉得自己很理性，甚至还期待他人也一样理性。我们看到他人为有严重缺陷的看法辩护时会生气，有时还会怒不可遏。但我们基本上不会想当然地认为那些与我们意见相左的人完全缺乏理性。我们恼火是因为发现这些人没有正确运用其理应具备的理性。（对我们而言）这么明显的事情他们竟然会不懂？

如果说理性是为发现事实真相所迫切需要的能力，那为什么拥有理性的人不能一直尽己所能去利用理性呢？毕竟，我们认为所有具备视力的人都看得见其他人能看见的东西。如果给几个人看一棵树或是一次日落，你理所当然地认为他们都会看到树和落日。然而，如果让几个人对不同问题进行思考，从逻辑性题目到社会性事件，若是他们最后得出了一样的结论，你反而可能会惊讶。若理性像知觉一样，能让我们弄清楚事物原本的样子，那就令人费解了。

对此，笛卡儿有自己的解释：“我们的意见如此多样并不是因为其中某些人更为理性，而只是因为我们指导自己思想的方式不同而已，而且我们也不会考虑相同的事情……最有才智的人能做最好的事，也能做最坏的事。”[注]

然而，这也只不过是谜团的重述罢了。我们难道不应该用理性来指导自己的思考吗？难道不该先用理性让自己免遭智力缺陷的荼毒吗？

在理性的众多支持者中，笛卡儿是最具说服力的。但理性常常也有许多激动易怒的贬低者，它的效力也曾遭受质疑，还因其傲慢自大而受到公开的谴责。宗教改革者马丁·路德（Martin Luther）则尤其尖锐刻薄："理性就其本质而言是一个恶劣的娼妓。不过她应当不会来荼毒我，只要我能够抵抗她。啊，可她又是如此美丽而闪闪发光……你一定要阻止理性，不要跟随她看似美丽的思考与计划。应向她的脸上投掷污垢，让她变得面目可憎。"㊟

平心而论，与这些断章取义的引言所暗示的内容相比，笛卡儿也好，路德也好，他们关于理性的观点要丰富得多，也巧妙得多，因此，也并不是那么激烈地反对理性。路德对于理性的猛烈抨击主要针对理性在宗教信仰方面的情况。在不同环境下，路德描述的理性就会变得更为传统，就像是"所有艺术品、药物、法律法规、人类现世所拥有的任何智慧、力量、美德及光荣的创造者与指导者"，以及"人与动物、其他事物相区分的关键差异"。㊟笛卡儿基于自身的坚定信仰，或出于谨慎的考量，避免了在理性的光辉下以批判性视角检验宗教信仰的忠诚度。

但理性若受到审判，原被告双方都能够做出非凡而卓越的陈述说明。被告方将引用笛卡儿、亚里士多德（Aristotle）、康德（Kant）或波普尔（Popper）的言论，提出人类犯错是因为思考得还不够。而原告方将引用路德、大卫·休谟（David Hume）、克尔凯郭尔（Kierkegaard）或福柯（Foucault）的言论，认为人类犯错是因为思考得太多。

与此同时，原被告双方也能够拿出令人信服的陈述来支撑自身观点。

埃拉托色尼与智能炸弹客㊟

你怀疑理性的力量吗？那就看看科学界吧，被告方将会大声呼喊。科学家通过极富洞察力的推理，发现了许多原本根本不可能接触到的隐藏事实与深奥见解。22个世纪之前，埃及亚历山大（Alexandria）港的古代世界最大的图书馆馆长埃拉托色尼（Eratosfhenes，公元前276—公元前195年）对于地球周长的测量成为既简单又极具说服力的例证；现代科学为理性的力量提供了数不胜数的例子，但没有一个能够胜过埃拉托色尼的例证。(注)

那时，人们已普遍认为地球是球形而非平面的，这很好地解释了海面上地平线的弯曲度及日月星辰显而易见的运动情况。但正如俗话所说，这仍然只是“纸上谈兵”而已。没有人真正地环球旅行过，更不用说像现在的宇航员一样从太空中观察地球了。那么，怎样才能完成对地球周长的测量呢？

埃拉托色尼听说，正午的阳光会在每年某天直射到远处城镇赛伊尼［Syene，现在的阿斯旺（Aswan）］的水井底部。他明白这就意味着太阳有时会运行至穹顶——赛伊尼城的垂直正上方，因此赛伊尼一定在北回归线上，而那天就一定是夏至（我们现在说的6月21日）。因此埃拉托色尼认为，赛伊尼城跟亚历山大港一样位于子午线上，朝向正南方。他知道商队从亚历山大港出发到达赛伊尼城需要多长时间，在此基础上就可以估出两城间的距离是5 014斯塔德［古希腊长度单位，约合607英尺（约185 014米）］。

夏至正午时分，当太阳位于赛伊尼城正上方时，太阳又在更北边的城市——亚历山大港偏南多少角度呢？埃拉托色尼用图书馆前方的一座方尖石塔（根据传说大致如此）来测量那一时刻投射下来的阴影。他认为太阳光线是以偏南7.2°的角度照射到方尖石塔上的。埃拉托色尼知道太阳离地球很远，远到它足以能让所有到达地球的光线呈平行状态，那么太阳光线与亚历山大港垂直面之间的角度，亚历山大港和赛伊尼城之间垂直面的角度，这两者是一致的，两条线在地球的正中心相交（见图

1.1）。也就是说，7.2°这个角度同时也是亚历山大港与赛伊尼城之间的纬度差异。现在，埃拉托色尼就获取了所有必需的信息。既然7.2°是360°的1/50，那埃拉托色尼就能够通过把亚历山大港与赛伊尼城之间的距离乘以50倍，来计算出地球的周长。演算结果是252　000斯塔德，比现代地球测量结果——24 859英里（约为40 008千米）少了1%。㊟

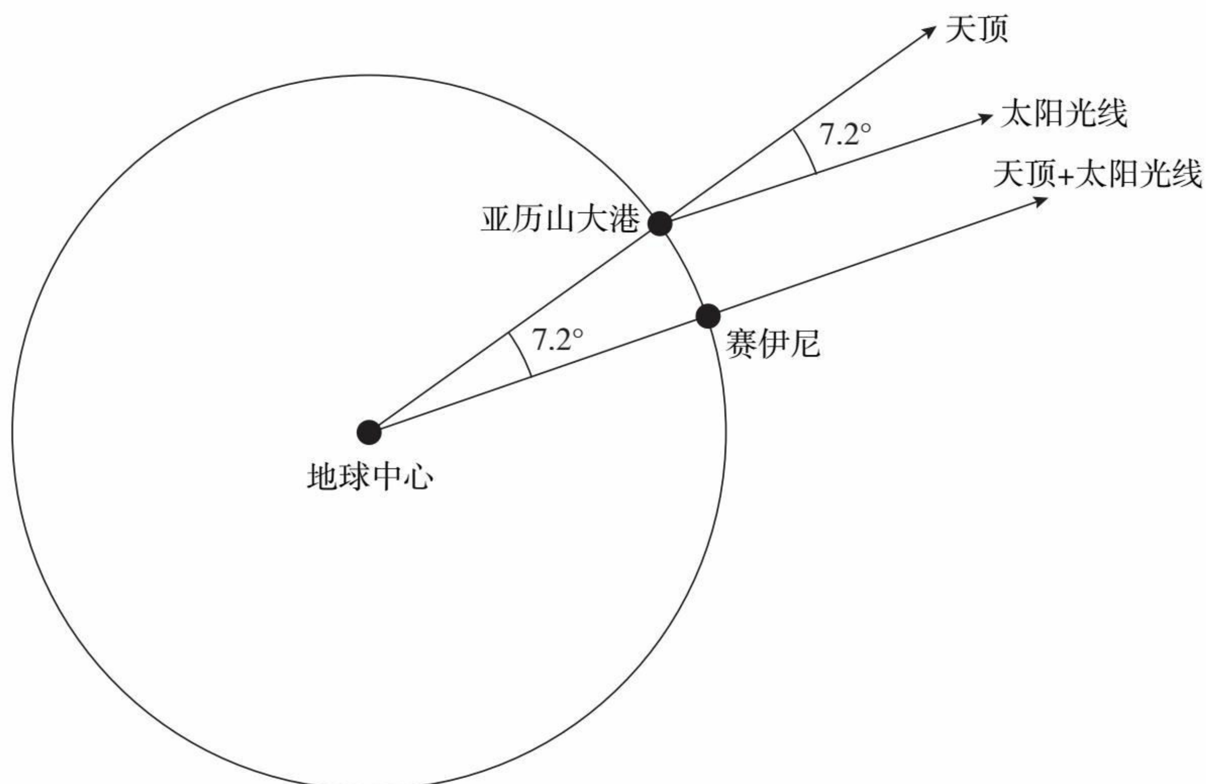


图1.1　埃拉托色尼是如何测量出地球周长的

埃拉托色尼抓住了两个证据间的关联性，可这两个证据从表面上看并无关联（商队的行进节奏、直射井底的光照及方尖石塔的阴影），他也抓住了假设之间的联系（地球是球形、日地距离）以及有关角度和平行线的简单几何学思想。他利用所有的线索，测出了一个他能够想象却无法看见也无法实测的圆周长。是什么让他的测算真实可信呢——难道不真实、不可信吗——这个测算结果纯粹是人类理性的产物。

原告方说，这样一项卓越的成就多么能够说明问题啊，理性的被告方竟然会选择站在另一边！这项成就确实超乎寻常，而且这也是2　000

多年后的今天它仍被人们铭记的原因。普通的推理并不会让我们走多远，只会将我们引向错误的方向。并非所有人都能效仿埃拉托色尼的模型，即便出色地运用理性，许多思想家还是严重地误入了歧途。出版社、报社以及科学杂志每天都会收到未来哲学家、科学家、改革家经过认真推理的无稽之谈，这些人总是无法得到出版社的认可，因此又去万维网尝试。然而，其中一些人的推理不仅从理论上讲是荒谬无理的，从实践上讲也是如此，施行这些推理结果只会带来昭彰的恶名，甚至是骂名。在这个节骨眼儿上，原告方可能还会提出泰德·卡钦斯基的例子。

作为一个年轻人，卡钦斯基无疑非常擅长推理。他于1958年进入哈佛大学，那时他才16岁。为了完成自己在密歇根大学的博士学位论文，卡钦斯基解决了一个困扰其教授多年的数学问题，为此，伯克利大学向他发出了聘用邀请函。然而两年后，他放弃了数学，也退出了学术界，隐居在蒙大拿州一个简陋的小屋里。在这个简陋的小屋里，卡钦斯基孜孜不倦地读着社会科学及政治类书籍。他的阅读与写作都集中于那些认为现代科技具有破坏性的主题上。科技进步会给自然环境及人类尊严带来灾难，这种观念在西方社会并不罕见，但卡钦斯基更加激进，他认为，只有使现代文明轰然倒塌的暴力革命才能够阻止这些以及更为可怕的灾难。

为了促成这场革命，卡钦斯基从1978年开始向大学、企业界及个人寄送炸弹，并造成三人死亡，无数人受伤。1995年，他写了一篇长长的宣言，并设法刊登在了《纽约时报》（*The New York Times*）和《华盛顿邮报》（*The Washington Post*）上，承诺事成之后会“停止恐怖主义”。美国联邦调查局（FBI）称呼他为“恐怖炸弹客”，最终于1996年将其逮捕。目前卡钦斯基正在美国科罗拉多州的监狱中服无期徒刑，毫无假释的可能。在那里，他仍在不停地阅读与写作。

在这个极富才华的年轻数学家身上到底发生了什么呢？卡钦斯基被自身的理性祸害了吗？他由此变成了媒体描述的“胡言乱语的疯子”了

吗？卡钦斯基的家人曾为保护他而做了些安排，试图以精神失常为理由进行申辩。理性的被告方一定会毫不犹豫地同意：非理性才是罪魁祸首。然而，卡钦斯基在拘留期间不太可能患上任何重度精神疾病，毕竟他仍是头脑灵活、口齿伶俐、知识极其渊博的人。原告方仍会坚持认为，卡钦斯基的所作所为并不能归咎于存在缺陷的推理。为了认识到这一点是正确的，你需要看一看这个“恐怖炸弹客”写的宣言。

> 工业革命及其后果对人类来说一直都是一场灾难……它们打破了社会的平衡，使生活毫无愉悦可言，使人类毫无尊严可言……还使自然界遭到毁灭性的破坏。科技继续进步只会让形势愈加糟糕……工业科技体系有可能存活，也可能轰然崩溃……如果体系存活了下来，那么后果就无法避免：没有任何办法可以改革或是修正这个体系，来让人们的尊严与自主权免遭剥夺。如果这个体系崩溃了，还是会给人类带来痛苦。因为这个体系发展得越大，它崩溃之后的结果也就越具有灾难性，所以如果它注定要崩溃，那还是尽早崩溃为好。因此，我们提倡要对工业体系进行一场革命。注

这确实是一篇布局极佳的议论文。大多数人并不同意文中所提的前提，即科技进步显然是一场灾难，但事实上，许多受人尊敬的哲学家、社会理论家都曾支持相似的观点。理性的原告方会认为，使卡钦斯基变得执迷不悟、一意孤行的，就是他将这个激进而悲观的看法当成了一种必然结果，并且据此有所行动。卡钦斯基的一位传记作者说道：“简而言之，卡钦斯基变成了冷血杀手，这并非与其才智无关，而是由其才智导致的。”注

所以，理性的被告方会反驳，原告方想让你相信，“恐怖炸弹客”泰德·卡钦斯基的问题在于推理得太多。他的宣言确实要比许多政治论述的推理更为严密。然而使他臭名昭著的并不是他的思想，而是他的罪行。在其作品中，没有任何一处（即便是布局合理的论文开头）表明向

无能为力的几个学者（即他以前的同事）寄送炸弹会触发“对工业体系的革命”。当你得知过度依赖推理会让人得出荒唐怪诞、令人憎恶的结论时，审视该证据便会发现理性的过失之处：某些前提并没有经过适当的检验，论证中某些至关重要的步骤竟然缺失了。记住：逻辑证明只要有短板，就不可能变强。

原告方的专家证人

既然历史例证无论多么富有吸引力都没有足够的效力来论证自己，那么理性的原被告双方就会转而求助自己的专家证人。双方要找到支持自己说法的心理学家都不是难事。研究理性的专家之间也没有达成一致。事实上，这些专家身陷其中的争论本身，就其激烈程度而言就足以被描述为“合理性战争”。有人期望专家都是明智的推理者，这样，专家之间意见不一致的情况就显得尤为可笑：对推理问题进行复杂推理，并不会让双方对问题达成一致意见。

理性的原告方可能会相当得意。关于推理的实验心理学自20世纪60年代以来发展迅猛，充分利用了各种巧妙的实验。就算再有名气的难题，原则上用一点点简单的推理也可以解决。但这些实验中大部分参与者都很自信地给出了错误的答案，仿佛他们就是某种“认知错觉”的受害者一样。一直以来，这些结果在理性战争中都用于支持人类理性存在严重缺陷的论断。理性的捍卫方抗议道，这种实验不过是人为捏造出来、用来误导大众罢了。这些实验的目的好像就是欺骗理智之士，让他们看起来傻乎乎的，而非为了理解理性的一般流程。想出这些实验的心理学家坚持认为，正如视错觉（又称视觉错误，visual illusion）揭示了普通的、精准的视觉的重要特性一样，认知错觉也揭示了一般性推理的重要特性。㊟然而，哲学家、科学作家及记者关注的是评估人类理性的这项研究看似不太乐观的影响，甚至还夸大了其悲观意味。

做算术时，不论你加减的数字是顾客、树还是星星的数量，不论这些数字代表的集合数量具有代表性还是人们闻所未闻，都没有关系，你只是在处理这些数字时运用算术规则而已，其他一切在你眼里都不重要。同样，如果你想当然地认为推理就是将逻辑应用于给定前提，以得出符合这些前提的结论，那么就没什么能干扰你了。然而，现在有充分的证据显示，背景知识与期望值确实会在推理过程中造成干扰。许多人认为，这就是产生糟糕推理的主要原因。

下面是一个经典例子。(注)1980年7月，比约恩·波格（Björn Borg）获得了温网五连胜，那时他被誉为有史以来最厉害的网球运动员之一。同年10月，在北美工作的以色列心理学家丹尼尔·卡尼曼与阿莫斯·特沃斯基（Amos Tversky）（他们很快就会闻名于世）向一群来自俄勒冈大学的大学生提出了如下问题。

> 假设比约恩·波格于1981年打进了温网决赛，请将下列结果进行排序（可能性由大到小）。
>
> 1. 波格会赢得比赛。
> 2. 波格会输掉第一局。
> 3. 波格会输掉第一局，但会赢得整场比赛。
> 4. 波格会赢得第一局，但会输掉整场比赛。

72%的学生都认为第三个结果的可能性更高，而非第二个结果。这有什么值得特别注意的地方呢？假若你有两个命题（如“波格会输掉第一局”和“波格会赢得整场比赛”），那么将这两个命题连接起来看（“波格会输掉第一局，但会赢得整场比赛”），实现的可能性就会比任何一个单独的命题更高。波格不可能在输掉第一局又赢了整场比赛的情况下没有输掉第一局，但他有可能在输掉第一局的情况下没有赢得整场比赛。认识不到这一点，就是所谓“合取谬误”（conjunction fallacy）的一

个实例。更抽象一点就是，分别用字母P和Q代表两个命题。当P且Q是真命题时，P与Q一定都是真命题；当P可能是真命题或Q可能是真命题时，“P且Q”就是假命题。因此，对于任何P命题和Q命题，明显错误的一点是认为P且Q要比单独的P命题或Q命题实现的可能性更高。

卡尼曼与特沃斯基设计的许多问题促使人们去犯“合取谬误”等严重错误。的确，正如他们自身表明的那样，如果问题涉及的对象不是温网中的比约恩·波格，而是普通比赛中的无名选手，人们就不会产生这种谬误了。他们就会正确地认为，单一事件发生的可能性比那一事件与另一事件相结合的可能性更大。然而，人们究竟为何会在一件有关无名选手的事上，比在一件有关著名冠军的事上要推理得更出色呢？

下面是我们自身研究中的又一个例子，阐明了你建构逻辑性问题的方式如何显著地影响人们的表现。(注)我们为人们呈现了逻辑上被称为“鸽巢问题”（pigeonhole problem）的下列版本。

> 丹顿小镇上有22个农民。所有农民都至少有一头奶牛。没有农民拥有超过17头奶牛。请问丹顿小镇上至少有两位农民拥有恰好相同数量奶牛的可能性是多少？

只有30%的人给出了正确的答案，也就是说，有至少两位农民拥有相同数量奶牛的可能性是百分之百——而不只是可能。如果你没有看出这一点，那么这个问题的另一个版本或许能够帮到你。

对于另一个群体，我们给出了这个问题的另一个版本，这一个版本与上一个版本从逻辑角度来看是完全相等的。

> 丹顿小镇上有22个农民。健康检查员于今年2月1日到17日拜访了镇上所有农民。请问这名健康检查员恰好在同一天拜访了至少两位农民的可能性是多少？

这一次，70%的人都给出了正确的答案：百分之百。

正如波格问题与著名的奶牛问题表明的那样，大多数人是依靠背景知识来思考或表达这些问题的，其中并不涉及逻辑。背景知识足够多，他们就会答对，否则就会答错。原告方会争论道，这难道不就是人类理性有严重缺陷的明显证据吗？

被告方的专家证人

许多心理学家都将注意力集中在似乎能表明人类非理性的实验上，有一些心理学家却有不同的追求：找出让人类能够推理的心理机制与程序。

毫无疑问，某种简单的推理（从广义上说）一直在进行着，尤其是与他人交谈时。像“且”“或”“如果”这样的连接，以及动词“非”引出了最基本种类的逻辑推理。举个简单的例子：

> 杰克：我把我的雨伞借给你或苏珊了——我不记得到底是谁。
>
> 吉尔：嗯，你没有借给我！
>
> 杰克：呦，那我就是借给了苏珊。
>
> 吉尔：没错！

杰克也好，吉尔也好，都不需要学习逻辑便能得出雨伞借给了苏珊这个结论。注那完成这种推理又是依靠什么样的心理机制呢？根据其中一种理由，理解词汇“或”或“非”，则意味着头脑中存在以某种方式捕获了这些词语意思的逻辑规则。这些规则负责管理的推论，是由陈述中出现的这些“逻辑性”词语导出的。这里有一个“或”的例子（再次使用字母P和Q来代替任何两个命题）：

“或”规则：从“P或Q”和“非P”这两个前提中，可以推出Q。

一些心理学家［尤其是让·皮亚杰（Jean　Piaget）、马丁·布雷恩（Martin Braine）及兰斯·里普斯（Lance Rips）］注已经证明，我们进行逻辑演绎是借助这种逻辑规则或图式构成的“心智逻辑”。当杰克和吉尔推断，杰克把雨伞借给了苏珊时，他们就是应用了这种“或”规则。

根据一种替代性解释，“心智模型理论”（mental model theory）［由菲利普·约翰逊-莱尔德（Philip　Johnson-Laird）及鲁思·拜恩（Ruth Byrne）提出］注认为，不，我们头脑中并没有一种心智逻辑存在。我们有的只是一个程序，借助与情况示意图相媲美的模型，来表现并整合我们头脑中关于这些前提的内容。我们就是从这些模型中读出结论的。举例来说，在我们的模型中，杰克把雨伞借给了吉尔；而在替代性模型中，则借给了苏珊。如果杰克的说法是正确的，那么这两种心理模型不会同时正确，也不会同时错误。我们知道的是“借给了吉尔”模型是错的，那么就只剩下了“借给了苏珊”这个模型，因此得出结论，杰克是把雨伞借给了苏珊。

推理心理学主要是研究“心智逻辑”与“心智模型”这两种方式之间的冲突。你可能会好奇：这两者之间到底有什么区别？不过，这两者之间有许多共同点倒是真的。它们都认为人类拥有能够推出真正逻辑性结论的机制。它们都认为人类具备合理思考的能力，并且在这方面，它们与质疑人类理性的方式形成对照。

然而，“心智逻辑学家”与“心智建模师”绘制的这张图不是很乐观。两种方式都意识到，只有最简单的推理任务会让人犯错，导向毫无根据可言的结论。当推理任务开始变得复杂时，它同时也变得极其艰难，其表现便一塌糊涂。但是，是什么使推理任务变得复杂呢？两种理论的不同之处就在于此。心智逻辑学家认为，是必须实施的步骤数及必须遵循

的规则数使推理任务变得复杂。心智建模师则认为，是应该构造并整合以得出某特定结论的模型数使推理任务变得复杂。

理性的被告方希望这两个派别能淡化分歧，把注意力集中于它们共有的一条积极信息：人类具备逻辑推理的通用性机制。而原告方却会在受这两个派别的启发而完成的工作中，找到许多能够质疑这条积极信息的证据。

若有一种基本推理模式在日常推理及学界推理中最为普遍且最为重要，那便是众所周知的条件推理，即“如果……就……”形式的推理（见图1.2）。这种推理需要“如果P，就Q”形式的大前提。例如：

如果你弄丢了钥匙，就欠我们5美元。

如果纯银被加热到961℃，就会熔化。

如果有法院，就会有警局。

如果玛丽（Mary）要写论文，就会在图书馆学习到很晚。

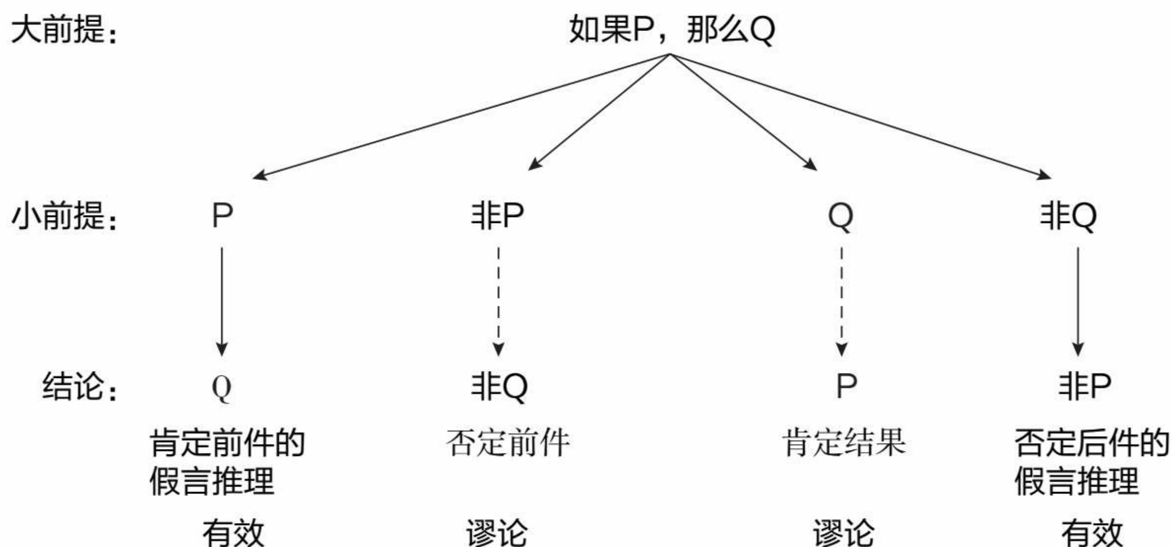


图1.2 假言推理的四种图示

这类陈述的第一部分由“如果”引出，是条件句的前件；第二部分

由“就”引出，是条件句的后件。要从条件语句中得出有用推论，需要用到第二个前提，而这个小前提要么肯定前件，要么否定后件。譬如：

如果有法院，就有警局。（大前提：条件语句）

有法院。（小前提：肯定前件）

所以有警局。（结论）

或者：

如果有法院，就会有警局。（大前提：条件语句）

没有警局。（小前提：否定后件）

所以没有法院。（结论）

这两种推论模式，一种是肯定前件（拉丁语名称是modus ponens），另一种则是否定后件（拉丁语名称是modus tollens），都是符合逻辑的：两个前提都为真时，其结论必然也为真。

但若把前件的否定（而非肯定）或后件的肯定（而非否定）当作小前提来使用，又如何呢？譬如：

如果有法院，就会有警局。（大前提：条件语句）

没有法院。（小前提：否定前件）

所以没有警局。（结论？）

或者：

如果有法院，就会有警局。（大前提：条件语句）

有警局。（小前提：肯定后件）

所以有法院。（结论？）

这两种推论模式（其小前提有“否定前件”与“肯定后件”两种）是站不住脚的，是谬论。即便两个前提都为真，结论也不一定为真——举例来说，可能有警局却没有法院。

原告方无疑会惊呼，这太简单了。如果被告方是正确的，难道人们不应该总是推出条件推理的两个有效结论，而从不会犯那两种错误吗？被告方的专家证人在问题十分简单的无数实验中表明，事情并不是这样的——远非如此。的确，几乎每个人都能从肯定前件中得出有效的演绎推论。这对于理性被告方而言是个好消息吗？余下的对于理性原告方倒是个好消息：平均只有三分之二的人能得出另一种有效的推论——否定后件式推论，并且大约一半人都犯了这两个错误。[注]还有更糟糕的……

她会在图书馆学习到很晚吗

在1989年的一项著名研究中，鲁思·拜恩表明，即便是有效的演绎推论——条件推理中唯一明显安全的那点儿逻辑性——也极易崩溃。[注]拜恩给参与者展示了以下两个前提：

大前提：如果玛丽要写论文，就会在图书馆学习到很晚。

小前提：她要写论文。

参与者很容易就能推断出：

结论：玛丽会在图书馆学习到很晚。

目前为止，一切都十分理想。然而，对另一群人，拜恩虽给出的是相同的问题，但附加了一个大前提：

大前提一：如果玛丽要写论文，就会在图书馆学习到很晚。

大前提二：如果图书馆一直开着，玛丽就会在图书馆学习到很晚。

小前提：她要写论文。

从严谨的逻辑学角度来看，第二个大前提没有任何关联。因此，如果人们富有逻辑性，就应该能像先前一样得出同样有效的演绎推论。但事实上，只有38%的人做到了。

拜恩并非想要证明人类是非理性的（心智建模师不相信这一点），而是想要证明心智逻辑学家对人类理性有错误的认识。若正如心智逻辑学家所言，人类有一套心理上的演绎推论规则，那么不论上下文语境如何，这一推论都应该是无意识的。参与者得到的指示是以前提为真，所以鉴于“如果玛丽要写论文，她就会在图书馆学习到很晚”与“玛丽要写论文”这两个前提，他们应该毫不犹豫就能推断出“她会在图书馆学习到很晚”。图书馆闭馆的可能性有多大呢？这又有什么关系呢？毕竟，就如你所知，即便图书馆闭馆，玛丽也许还是能在里面学习。心智逻辑学家会告诉你，不要那么想。这与此逻辑任务并无关联，正如气泡会爆炸的可能性与“两个气泡加三个气泡”的算术任务无关联一样。

鉴于拜恩的研究结论，心智逻辑学家没有意识到他们的方式是错误的吗？没有，他们并不一定要这么做。相反，他们给出了替代性解释。注譬如，人们可能会把拜恩给出的两个大前提整合为“如果玛丽要写论文，并且图书馆一直开着的话，她就会在图书馆学习到很晚”这一个前

提。毕竟这是了解这一处境较为实际的方式。如果人们就是这么解释两个大前提的话，那么小前提“她要写论文”就不足以推出有效的演绎推论，而且不论拜恩的研究结论多么有吸引力，都不能成为反对心智逻辑的证据。

在这次审判中有被告吗

看到心智逻辑学家与心智建模师——被告方的所有专家证人彼此争论不休，理性的原告方可能会觉得颇为有趣，但毫无疑问，此时此刻，陪审团可能变得没有耐心了。难道没有哪里出错吗？不是这些实验参与者的推理，而是心理学家的诉求出错了，难道不是吗？

实验主义者希望参与者认可前提为真，而不论这些前提合理与否，希望他们给出的只是从这些前提必然会推出的结论，希望他们完全忽视这种微乎其微的可能性，即忽视现实世界。若人们无法确定所给前提的逻辑内涵，许多心理学家便由此得出这些人的推理能力不足。但还有一种替代性解释，即人们得到的人为性指示是难以遵循的，甚至在很多情况下是不可能去遵循的。

这并不是因为人们不善于进行逻辑演绎，而是因为人们不善于将这些演绎与基于同一些前提的概率推论区分开来。可是，这能证明人类没有理性吗？这难道不能证明心理学家有不理性的诉求吗？

与视觉心理学进行比较会有所帮助。请看图1.3——爱德华·阿德尔森（Edward Adelson）设计的著名错觉实验。A、B两个方块哪个的灰度更浅？无疑，B的灰度更浅些——这不可能是错觉！但这确实是错觉。不论多么出人意料，A与B的灰度确定是一样的。

简要说来，这并非难以理解。你对方块表面颜色深浅的感知，是基于落在表面而被表面反射出去的光线比例，而非光线从哪一表面反射到

你眼中的数量。

这个“反射比”（就是上文说的光线比例）越高则表面越亮，反射比越低则表面越暗。

反射比=表面反射的光线/落到表面的光线

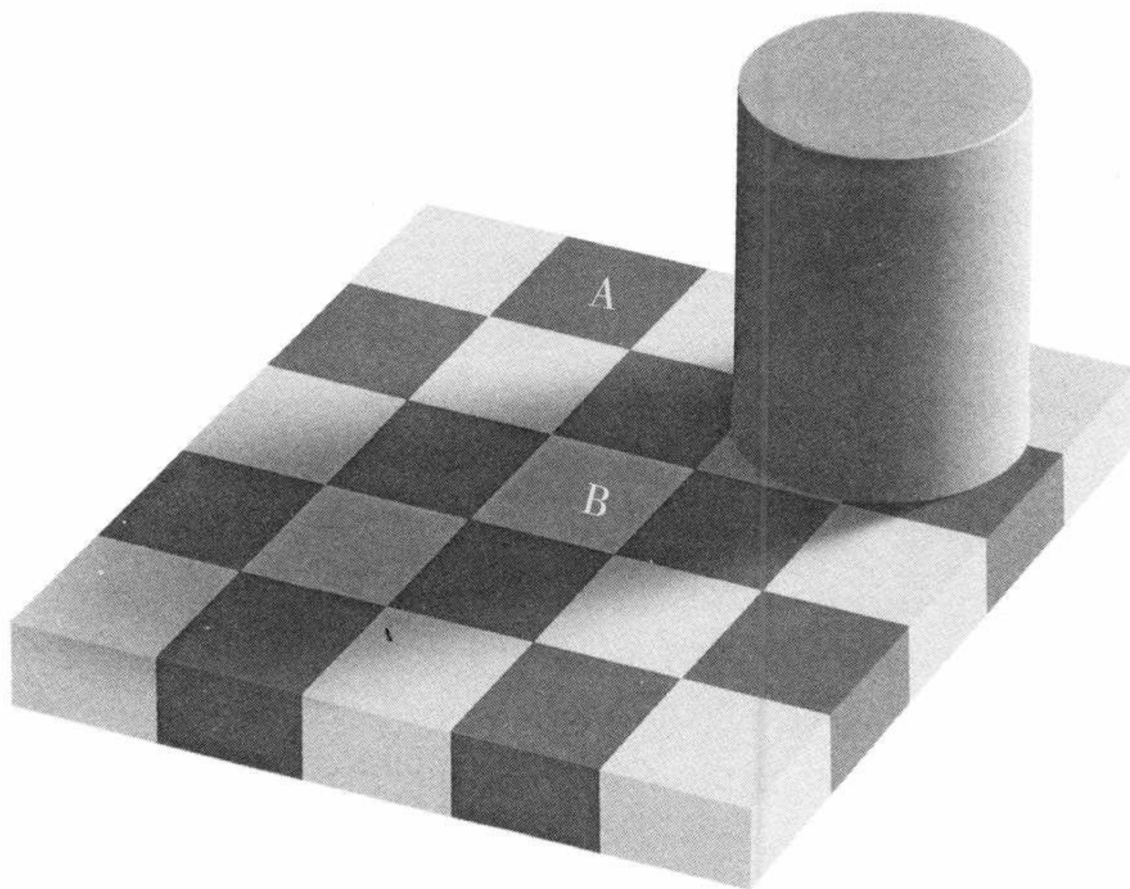


图1.3 阿德尔森的棋盘错觉

表面同为灰色，接收并反射或多或少的光线到你眼中，但若反射比仍然相同，那么你感知的会是同一灰度。可是，你的眼睛就只从这两个量中的其中一个（即反射到你眼中的光线）获取信息。那么，在只能感知其中一个量的情况下，大脑要如何锁定反射比（即两个量之间的比例）并估量表面的明暗呢？为此，必须要运用上下文信息及背景知识来推断出另一个相关的量，即落到表面的光线数量。

当你看图1.3时，你看到的是一张西洋棋棋盘的图片，其中有一部分处在一个圆柱形物体的阴影之下。

此外，你希望棋盘上的方块是明暗交错的。因此你有合理的理由去判断方块B——阴影中的亮色方块之一 ——要比方块A（直接接收光线的暗色方块之一）更亮。更确切地说，如果你真的是在看一个部分处于圆柱阴影之下的棋盘，而不仅仅是一张图片，你就会有充足的理由。你不能把这张图片只当作有不同灰度表面的二维图案，也不能无视所绘的立体图，因此便产生了错觉。

画家与平面设计师得学会克服把所有可能相关的信息进行整合这一自然倾向，而其他人则深受错觉之害。我们应该为这一错觉大吃一惊并感叹自己的视觉感知没有想象中那么好，甚至怀疑其可信性吗？恰恰相反！因为我们不仅能把视网膜刺激考虑进去，还能把对光物理学及物体结构的直觉领悟考虑进去，所以能分辨并理解所感知的事物。即便我们看到的是一幅画，而不是真正的物体，通常还是会对表征之物的性质感兴趣，而非这张画本身的物理性质。虽然纸上或屏幕上方块A图片与方块B图片的灰度相同，但是这张图所绘的棋盘中的方块A比方块B暗得多。视错觉让我们能较好地理解自身所处的三维环境，而且鉴于对图像的熟知，还让我们能较好地解读三维环境中的二维图片，这一错觉便是证据。

现在，回到那个可能会在图书馆学习到很晚的玛丽一例。通常情况下，我们解读条件语句时就是假设它们是相关联的。(注)因此，若给出第二个大前提“如果图书馆一直开着，玛丽就会在里面学习到很晚”，人们就会合理地假设这一前提也是相关的。要是这一前提相关的话，情况就一定是图书馆可能会闭馆，而这会让玛丽打消在图书馆学习到很晚的念头。不错，参与者得到的指示是要将“如果玛丽要写论文，就会在图书馆学习到很晚”视为是绝对真实的，但他们似乎并不这样认为。可是，无法遵循这样的指示，跟无法好好推理根本就不是一回事。将有意提供

给你的信息视为相关联的并非不理性，而恰恰是理性之举。

画家需要有耐心并勤加练习才能看出画布上本真的色彩，而非他人从整幅画的背景下感知的色彩。同样，学逻辑的学生也需要有耐心并勤加练习才能只注重前提中的逻辑术语，并无视乍一看可能相关联的上下文信息及背景知识。画家看到而我们没看到的，对画家而言才是真正有用的。逻辑学习者推出的结论对他们而言才是有用的。画家的视觉技能与逻辑学习者的推理技能在平常的生活中也有很大用处吗？如果并未想过要成为画家或是逻辑学家，我们是否就该为没有能够跟他们共享认知技能而感到错失了某些重要的东西呢？事实上，不必如此。

鲁思·拜恩实验中参与者保持理性的方式需要进一步研究，但相当明显的是，参与者确实保持了理性。他们没能够解决基本的逻辑性问题，并不意味着他们无法很好地进行推理，这恰恰与他们对现实生活问题的解决有关。这一边的逻辑与那一边的推理，这两者之间的关系一点儿也不简单、一点儿也不直接。

此时，法官、陪审团以及读者朋友可能已经对原被告双方做秀般的哗众取宠感到心力交瘁了。虽然这个审判构思是我们想出来的，但其中的论战（我们只做了几个简单的介绍）是非常真实的，已经持续了很长一段时间。对立双方的争论越来越尖锐，而争论的问题本身却越来越模糊不清。这场争论究竟是关于什么的？据称这种推理能力既让人类优于其他动物，又如此劣质，那它到底是什么呢？卡尼曼与特沃斯基的实验与心智逻辑学家和心智建模师的实验解决的是同一个问题吗？就此而言，他们讨论的推理与笛卡儿颂扬而路德贬低的理性是一回事吗？最后再设想一下，这场审判中有被告吗？若有，那它是理性本身，还是某种误以为真的虚拟物呢？真的有理性这回事吗？

---

1. Descartes 2006.

2. Descartes 2006，p.13。

3. Descartes 2006，p.18。

4. Descartes 2006，p.5。

5. Descartes 2006，p.5。

6. 在Galler 2007引用、翻译并讨论。

7. Luther 1536, LW 34:137 (nos.4-8).

8. 泰德·卡钦斯基（Ted Kaczynski），曾就读于哈佛大学，任教于伯克利大学，美国高智商反科技恐怖分子，反对工业文明与现代科技，多年以来一直向大学中高科技领域的研究者寄送炸弹，1998年被判终身监禁。——译者注

9. Nicastro（2008）对这个故事进行了很好的阐述。丹·斯珀伯在回答约翰·布罗克曼（John Brockman）的问题“什么才是你喜欢的深刻、优雅且美丽的解释？”时，对此有所引用。

10. 事实上，对这个赛跑场，暂时还没有一个公认价值：各国都有自己的尺寸版本。埃拉托色尼使用的赛跑场长度是多少呢？这上面存在一些争议，埃拉托色尼测出来的地球周长也就存在争议。但毫无争议的是，埃拉托色尼使用的方法具有创造性与合理性。

11. Kaczynski 2005.

12. Chase 2003, p.81.

13. 这个主题在Piattelli-Palmarini（1994）中论述得十分完善。

14. Tversky & Kahneman 1983.

15. Mercier,Politzer & Sperber，正在印刷中。

16. 举例参见：Prado et al.2015。

17. Braine & O'Brien 1998; Piaget & Inhelder 1967; Rips 1994.

18. 举例参见：Johnson-Laird & Byrne 1991。

19. 例如，可参见：Evans 1989。

20. Byrne 1989。在拜恩的实验中，在图书馆学习到很晚的那个人并没有名字，为了阐述方便，我们称其为“玛丽”。

21. 例如，Bonnefon & Hilton 2002; Politzer 2005。

22. 相关性在理解中起到的作用在Sperber & Wilson 1995中有所强调。

第二章
心理学家的艰难求索

有人认为，理性使人类与其他动物相区别，这种观念可以追溯到古希腊哲学家亚里士多德。注迄今为止，亚里士多德一直是西方思想史上最有影响力的思想家，长期以来被称为“大哲学家”，仿佛他是唯一堪当此名号的人。他成就众多，人们还认为他发现了逻辑科学。通过发现逻辑科学，他将理性作为将会永久存在的联合伙伴——看起来是这样。确实很少有联合体能够像逻辑与理性那样长久存在，但最近（大约100年内），这个联合体一直摇摇欲坠。

理性与逻辑？这很复杂

19世纪末以前，逻辑与推理研究虽非完全一样，但也不能说它们是同一组织的两个方面。逻辑被认为是对优秀的、正确的推理的描述。并非所有推理都是对的——正如我们看到的，远远不是如此——但就像人们一般认为的，所有的推理都应当、都旨在变得富有逻辑。糟糕的推理是试图富有逻辑却并非如此的推理（要不然就是装成富有逻辑的样子的诡辩术）。因此，即便我们经常毫无章法地表达自己，但就像是语法规定了语言的范围一样，逻辑也规定了推理的范围。

课本上的推理范例通常都始于一点简单逻辑，然后就戛然而止了，也不解释推理者脑海中到底在想什么。这些推理范例通常包含两个前提与一个结论。举例来说（我们在此引用亚里士多德运用所谓“直言三段

论”进行论述的极负盛名的例子中的一个）：

前提：1.所有人终有一死。

2.所有希腊人都是人。

结论：所有希腊人终有一死。

从命题“所有人终有一死”和“所有希腊人都是人”，就能逻辑性地推出“所有希腊人终有一死”。同样，从命题“杰克把雨伞借给了吉尔或者苏珊”和“他没有把雨伞借给吉尔”，就能逻辑性地推出他把雨伞借给了苏珊（这是“析取三段论”的一个例子）。亚里士多德逻辑学的成就之一便是采用这种清晰的例子来有效推理，并将之系统性整合。

忘了人类、死亡与希腊人。随便列出三个种类，称之为A、B、C。之后你就可以进行概括，认为具备下列概要形式的三段论都是有效的：

前提：1.所有的A都是B。

2.所有的C都是A。

结论：所有的C都是B。

忘了雨伞、杰克、吉尔和苏珊。随便列出两个命题，称为P和Q，你就又可以进行概括，认为任何具备下列模式（与我们在第一章中讨论过的“或”规则相对应）的三段论都是有效的：

前提：1.P或Q。

2.非P。

结论：则Q。

识别这种模式的意义何在？直觉认为，某些特定的演绎碰巧是有效的，如希腊人终有一死、杰克把雨伞借给了苏珊，这种识别的意义就在于把这种直觉变成正式陈述，不仅使这些特定的演绎有效，还包括同一形式的所有演绎。用如大写字母一般的任意符号来代替具体内容（一般认为与演绎无关）是亚里士多德独创的手段，这样做的话，你最终就会得到一个“逻辑形式”，其中只包含如“所有”“或”“非”这类与前提-结论关系相关的术语。

2 000多年间，学者觉得并无必要超越亚里士多德式逻辑。《纯粹理性批判》（*Critique of Pure Reason*）的作者康德在18世纪末竟认为，自亚里士多德以来，逻辑“一直无法取得更进一步的成就，由此看来，所有该达成的都已达成了”。㊟他这个观点简直错得离谱。

过去200年来，逻辑在范围和复杂程度上都已远远超过亚里士多德式起源，更是分出了许多不同的子领域和方法，甚至已经进入多元化发展阶段。与古典逻辑相比，现代演绎逻辑给更多种类的有效演绎提供了正式陈述。能做到这样，并不是因为罗列了演绎模式，而是因为用精巧的方式从第一原则获取了这种模式。然而，现代逻辑研究的许多演绎，哪怕是其中相对简单的演绎，既不属于普通人的其中一项技能，也非数学中的高级定理。的确，有些逻辑研究项目想要与心理学扯上关系，但是现代逻辑从整体上看并非如此。

推理的实证研究始于20世纪。㊟那时，许多逻辑学家把逻辑看作与数学密切相关的纯形式系统。现代逻辑学的奠基者——德国逻辑学家戈特洛布·弗雷格（Gottlob Frege）曾公开谴责“逻辑就是关于人类推理”的观念，认为这是“心理主义”谬误：逻辑与人类推理无关，就像算术与人类对数量的理解运用无关一样。这就是现在的主流观点。

虽然大多数逻辑学家都背弃了心理学，但是大多数推理心理学家还是要靠逻辑来定义其研究领域并将之分成各个子域，然后确定构成好的

或坏的推理的因素。直到最近，他们都几乎不曾想到这会发展成心理学中的“逻辑主义”谬误，就像逻辑学中的“心理主义”谬误一样。㊟

将推理看作一个“逻辑性”过程似乎确实颇为自然。当人们进行推理时，脑海中最先会出现一些想法，而这些想法也必须先出现，其他想法随后才会出现。把思绪的暂时性、随意性排序与演绎中命题的逻辑性排序相等同，这或许十分吸引人。虽然“结果”与“下列”这两个词语在逻辑上表现的是时间顺序，但其实并无时序关系。逻辑序列中的命题顺序既非真正的时间顺序，也非算术中“1，2，3，……”这样的正整数顺序。心理学过程会持续一段时间，并且需要一定的努力。逻辑序列则没有这些。

逻辑学中，“论证”一词描述了从前提推出结论的一个无时间性、抽象的命题序列。另外，平常的用法中，论证是某人在脑海中或在对话中拿出一个理由，或相继拿出好几个理由来为某个结论辩护时的产物。怎样做才能避免混淆呢？自从推理心理学集中讨论研究经典演绎论证，也称为“三段论”，这便成为我们会在批判性论述中用到的术语。另外，我们也会一直用到该词的普通含义、非术语义。

用来说服听众的一系列理由难道比不上由前提推出结论的逻辑序列吗？好吧，通常并非如此。通常情况下，你在论证时会先提出想要听众认同的结论（想一想正在为当事人进行无罪辩护的律师，或是想一想政治讨论），接着你会给出支撑这一结论的理由。人们仍普遍认为，大多数普通推理论证，即便不是所有的，无论怎样都是论证，必须与三段论对应；若这种对应并非显而易见，则必暗含其中；为图简练，某些前提必定已经省去。据此，大多数普通论证都是“省略三段论”，即缩减了的三段论。我们认为，这只是陈规罢了，太过理所当然，以至于几乎不用花任何工夫，单凭经验就能证明其合理性。

逻辑与推理心理学彼此如此接近，侧重方向却有所不同。尽管如

此，它们还是有许多理念是共通的，但其实它们相同的只是标签而已，这两个词在各个学科下呈现的是不同的含义，让人分不清。(注)“论证”并非唯一一个可以同时描述抽象逻辑性事物和具体心理学现象的词汇。许多像“推论”“前提”“结论”“有效”“合理”这样的词，都是从一个领域借用到另一个领域，在两边同时使用的过程中却很少有人关注它们用途的不一致。甚至“推理”这个词也被逻辑学家用来讨论三段论、逻辑推导以及证据，“合逻辑的”一词也常被用作心理学术语（如“注意逻辑性！”）。我们会努力避免这种多义造成的错误。

有些面包师是运动员

推理心理学对逻辑学的单恋造成的后果比较严重。举例来说，许多著名的心理学家选择探究人们如何运用亚里士多德的直言三段论。为什么？因为2000多年来，这些三段论一直是经典逻辑的重心，所以它们无疑会在心理学中发挥重要作用。

当所有能钻的牛角尖都钻了以后，出现了直言三段论的256种可能形式（若包含计数变量就是这个数的两倍），且每种都可以通过实验加以验证。为此，许多研究者研究多年。可惜的是，在这些实验中成千上万的参与者需要解决的系列问题也十分无聊而啰唆，这些问题被以下列方式相继提出。

有些面包师是运动员。

面包师都不是皮划艇运动员。

如果可以得出结论的话，会是什么结论呢？

在这256种三段论形式中，只有24种在逻辑上是成立的。目前尚不

清楚，在这些成立的和不成立的三段论中，哪些形式会出现在人类的实际推理中，出现的频次又如何。

当然，在科学中，对于边缘现象或实际上并不重要的现象的研究可以具备重大的科学相关性［想一想普通的果蝇，也称“黑腹果蝇”（Drosophila melanogaster），及其在现代生物学中的地位］，但这不是一个很好的例子。经过半个世纪的深入研究之后，桑吉特·科姆拉尼（Sangeet Khemlani）和菲利普·约翰逊-莱尔德于2012年发表了一篇评论文章，确定了三段论推理中12种矛盾理论，他们表示，没有一种“表述充分”。他们补充道：“任何科学领域存在这12种矛盾理论，都是一场小小的灾难。”注

由于约翰逊-莱尔德本人对三段论的心智模型表述是12种矛盾理论中最具有影响力的，因此，这项控诉变得更加有力，甚至更加尖锐。

支持运用不同推理方式［心理逻辑、心理模型、贝叶斯推断（Bayesian inference）等］的人运用了直言三段论研究来证明自己的方式才是最好的，可是也同样证明只有已改变信念的人才觉得三段论有说服力。

除了心理学家之外，另一批学者认为经典三段论仍具有高度相关性，这些人便是神学家，这些人从中世纪以来就一直在教授并运用“三段论法”。譬如，家喻户晓的神父、教皇本尼迪克特十六世（Pope Benedict X VI）的御用神学家伍杰·盖尔蒂赫（Wojciech Giertych）便是这样解释女性为何不适于神职的：“就哲学定义与逻辑三段论而言，男性更可能去信仰上帝。”注不相信吗？完整的亚里士多德式三段论与心理学的关联性同样神秘，对此，我们忍不住要予以讥讽。

“永远不要做你知道为何而做的实验！”

只有极少数推理心理学家（若有的话）在该领域的影响力大过彼得·沃森（Peter　Wason）。约翰逊-莱尔德说：“沃森做研究的方式颇为古怪，譬如永远不要做你知道为何而做的实验！”[注]

1966年，沃森采用了四卡选择任务这一新实验设计，该实验设计成为本学科研究的主要工具和重心，至今仍是如此。只有极少数研究者抵抗住了其吸引力，这些人认为，说推理心理学在很大程度上已变成了沃森任务心理学并非完全错误，只是有些夸大其词、有失公允罢了。如果沃森在构思这一实验时并不知道该实验会有何作用，那么人们肯定会认为，最后这一实验会误打误撞取得惊人的成功。

沃森实验的步骤如下。如图2.1所示，实验者会告诉你：“在你面前有四张卡片。每张卡片一面写有字母，另一面写有数字。两张（写有E与K的）卡片字母面朝上，另两张（写有2与7的）卡片数字面朝上。”

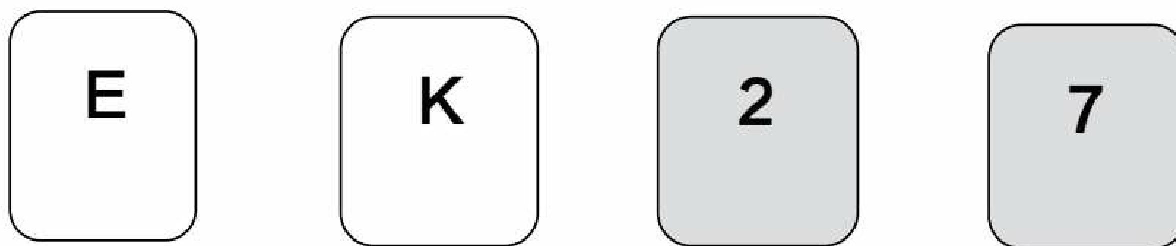


图2.1　沃森选择任务的四张卡片

“你的任务就是回答下列问题：必须把这四张卡片中的哪一张翻过来，才能弄清有关这四张卡片的规则，从而判断‘如果一张卡片一面写着E，则其另一面写着2’这句话的真假？”

你会选择哪张卡片呢？

该实验的构造源自经典逻辑学中的标准推论类型，即我们在第一章中见过的假言三段论。图1.2列出了假言三段论的四种模式，图2.2则展示了该选择任务是如何建立在四种模式的基础之上的。

该选择任务的“规则”是“如果P，则Q”（我们的例子是“如果一张卡片一面写着E，则其另一面写着2”）形式的假言三段论的大前提。每张卡片可以提供一个小前提（在我们的例子中则表现为，写有E的卡片代表小前提P，即卡片一面写着E；写有K的卡片代表“非P”，即卡片上没有写着E；写有2代表Q，即卡片的另一面写着2；写有7的卡片代表“非Q”，即卡片上没有写着2）。你可能还记得，这些小前提只有两个，即P和非Q可以促成有效演绎（分别称为肯定前件推理和否定后件推理）；若试图从另两个小前提，即“非P”和Q来进行相似的演绎，则会产生“否认前件”和“肯定后件”的谬误。

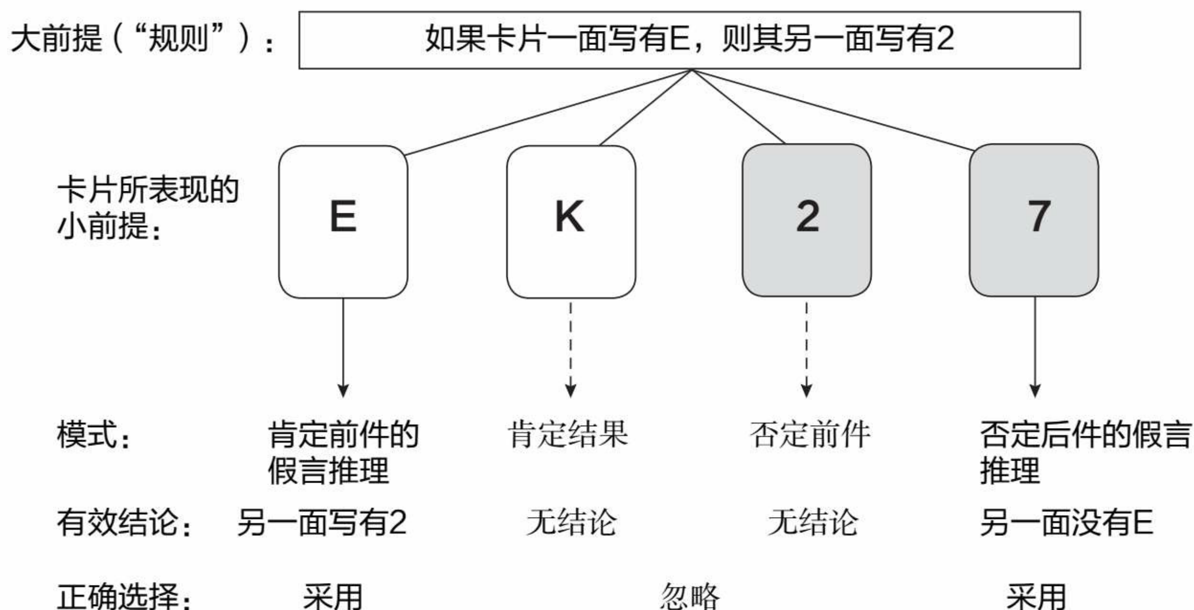


图2.2 选择任务及假言推理的四种模式

因此，正确的做法就是只选择写有E和7的卡片。这条规则蕴含了对这两张卡片另一面是什么的预测，而且把这两张卡片翻过来就能检验该预测。要是两张卡片的另一面都与预测不符，那么该规则就是伪造的。换句话说，这条规则就不蕴含任何对写有K和2的卡片另一面是什么的预测；因此规则与卡片并不相关。值得一提的是，与通常的直觉相反，翻开写有2的卡片并没有用。假设卡片的另一面写的不是E，又如何呢？规则所说的无非就是，卡片一面写着E，另一面就一定写着2；但规则并

没有说卡片一面写着2，另一面就只会写着E。

若你选了写有E和7的卡片，那么祝贺你！若你选了另外两张卡片，也不要太沮丧。无论怎样，只有约10%的参与者做出了正确的选择。

心理学家一旦开始用沃森任务来实验，要停下来就很难了。很多人都陷入其中无法自拔。这是因为，该选择任务的设计适用于各种变化。你可以更改指示，修改卡片内容，或构思各种语境。接下来你可以观察到发生了什么，特别要留意一下，与原版沃森任务相比，变化后的版本中做出正确选择的参与者是否变多了。若果真如此，写篇文章吧。若并非如此，就再试试。况且，不仅心理学家，哲学家、学生，甚至你的室友或表亲也能很容易就提出猜想，对参与者为何做此回应进行解释，并建议一些新的变动。选择任务证明了交谈的永恒主题，众人——不论是专业人士还是业余爱好者，都可以试试。

当然，从理论上来讲，该选择任务的成功之处就在于，能够卓有成效地提供关键证据并帮助回答根本问题，就像显微镜对生物学的作用一样（此类比我们曾有耳闻）。该选择任务有促成任何一项理论性突破吗？没有，只要人们声称实验证据为某一真的理论主张提供了关键性支撑，就会有替代性解释被提出来。因此，用该任务做的众多研究把解释任务本身定为其目标，(注)而人类推理心理学只是作为昏暗的背景来烘托对实验精彩纷呈的辩论。

早期研究的许多内容都旨在改善人们在选择任务中的糟糕表现。训练会有帮助吗？帮助不大。那反馈呢？几乎没有。改变规则的措辞呢？不妨试试。对良好表现给予物质奖励呢？算了吧。那么，变动选择任务的非逻辑内容（用更有趣的东西来替换数字和字母）会让人们有更好的表现吗？是的，有时候可以，但解释起来就难了。

那么在众说纷纭中，有哪些好的方面呢？有些发现不仅与理解任务相关，还与理解心智相关，这些发现通常是偶然的，而非先预测后证

实。选择任务的故事主要阐明的是，优秀的科学家如何能不断探索一个接一个的死胡同。

讽刺的是，50年来对此任务的研究最重要的发现竟是，人们根本就不会运用推理来解决意在展现其如何推理的任务。

20世纪70年代早期，乔纳森·埃文斯（Jonathan Evans）对标准选择任务稍加变动，然后进行检验，结果令人费解。[注]拿运用这一规则的一般性问题举个例子，“如果卡片一面写着E，则其另一面写着2”，共有四张写着E、K、2、7的卡片。正如我们所见，只有约10%的人选择了写着E和7的卡片，虽然这才是符合逻辑的正确做法。现在只是在规则中加了“不是”二字，像这样：“如果卡片一面写着E，则其另一面写的不是2。”给相同的卡片，问相同的问题。现在，绝大多数的参与者都给出了正确答案。

不要妄下结论，认为否定性规则能让参与者成为优秀的逻辑推理者。事实上，在两种情况下（规则被否定或未被否定），大多数参与者还是会选择写有E和2的卡片，仿佛否定的存在没有丝毫影响。凑巧的是，该选择在标准情况下是错误的，但在规则被否定的情况下是正确的。（怎么会这样？因为，肯定性规则没有预测写有2的卡片另一面的字母是什么，但否定性规则做了预测：写有2的卡片另一面若写着E的话，就会证明否定性规则是错误的。所以对否定性规则来说，就应选择写有2的卡片。）

埃文斯认为，这表明人们对沃森任务的回答并非基于逻辑推理，而是基于对相关性的直觉：他们翻开直觉上看起来相关的卡片。为什么写有E和2的卡片看起来有直觉上的关联呢？埃文斯解释道，因为二者在规则当中有所提及，而没提到其他字母和数字，这几乎就是全部的原因了。[注]

沃森选择任务冗长且令人费解的故事，很好地解释了人类推理心理

学如何以及为何最终渐渐背离当初对经典逻辑学的痴迷，并迎接崭新的挑战。

双重过程？

看看对选择任务的研究，再更广泛地看看推理实验心理学，你就会看到心理学家在尽力做到全面详尽。再看看研究工作些微的进步以及毫无指望的整体形势，就会使人更加困惑、更加丧失信心——而在这一时期，认知研究取得了显著的进步。许多领域（如视力、婴儿心理学以及社会认知等）都有重大发现、新颖的实验方式及明显的理论突破，且发展速度越来越快。每个月，学术期刊都会刊印振奋人心的新成果。虽然辩论激烈，但研究员有清晰的共同目标且强烈希望共同进步，在推理心理学中根本不是这么回事。的确，有些思想流派声称取得了重大突破，但不论理由好坏，这些所谓的“突破”都未受到广泛认可。

尽管如此，如果做个调查，问及推理心理学家该领域近来最重要的理论性成就，大多数人（尽管少数人强烈反对）都会说是“双重过程理论”：认为在推论中，或者更宽泛地说，在人类心理学中，包含两种截然不同的基本过程类型。

双重过程理论的一个基本洞见就是，认为人们为解决推理任务所做的许多工作其实根本就不是推理，而是其他类型的过程，比推理更迅速、更自动化、更无意识，也更少受制于规则。在过去20年中，此方式已拓展为众多不同版本。我们担心，对“系统一”和“系统二”的讨论正变得与对“右脑”和“左脑”的讨论几乎一样普遍、一样空洞无意义。

其实，1975—1976年，乔纳森·埃文斯与彼得·沃森在发表的多篇文章中阐明了双重过程理论早期的核心内容，但很快就被人遗忘了。正如我们所见，埃文斯只是在选择任务的规则中加了“不是”二字，就论证了

人们在做出选择时实际并没有用到推理。他们只是选择了自己直觉上认为相关的卡片（碰巧对原规则来说是错误选择，对否定性规则来说是正确选择）。这样的话，该选择就是基于类型1的直觉过程。

埃文斯和沃森重做了实验，这一次却要求人们对自己的选择进行解释，参与者确实进行了推理，这一点毫无疑问。只不过他们并不是为了解决问题才进行推理（因为该问题他们已经凭直觉解决了），而是为了佐证其直觉答案。当参与者的答案碰巧从逻辑上看是正确的时（通常会在否定性规则下成立），他们便能给出符合逻辑的正当理由。当他们的答案碰巧不正确时，他们会用同等的自信给出毫无逻辑的理由。这样看来，有意识的推理（类型2过程）就只是为发生在实际推理之前的选择进行“合理化”。

在对双重过程方式的第一次概述中有三点值得注意。18世纪苏格兰哲学家大卫·休谟及19世纪美国哲学家威廉·詹姆斯（William James）着重指出，第一点是再度运用众多对比中的一种，这种对比发生在两种推论模式之间，一种无意识且不费力气，另一种（即合理化推理）则相反，有意识且费力气。第二点更新奇，认为人们可能且经常在两种模式下处理相同的推论任务。举例来说，在选择任务中，大多数参与者既凭直觉选择了卡片，又对自己的选择做出了合理化解释。第三点最有煽动性，认为类型2思考过程通常只是合理化类型1直觉过程所达成的结论。这一观念太贬低推理的特定作用了，以至于人们对埃文斯与沃森的双重过程方法持保留意见或怀疑态度。注

在接下来的20年里，这个早期的关于推理的双重过程方法并未经常被人提及，更不要说讨论了。当它再次出现在人们的视野中时，已经没有了早期理论中的多余部分；不复存在的那部分观念认为，推理只是将由其他方式得出的结论进行合理化罢了。因此，埃文斯与哲学家大卫·奥弗（David Over）于1996年合著了一本书——《合理化与推理》（*Rationality and Reasoning*）注。在此书中，他们在提倡“思考的双重

过程理论”的同时，将类型1过程也合理化了，并且把类型2过程“从纯粹的合理化作用变成表现逻辑成分的基础”。此外，最初的假设认为，这两种过程类型遵循严格的次序，先是无意识的决定，然后是有意识的合理化，但该假设被1976年顺便建议的一个替代性说法取而代之，即这两种过程类型是相互作用的。尽管埃文斯和沃森提出的双重过程理论的早期版本削弱了人们的合理化主张，但其后期版本证明了这些主张的正确性，据说甚至还对此有所扩展。

显然，时机成熟了。在同一年，也就是1996年，美国心理学家史蒂文·斯洛曼（Steven Sloman）发表了“两种推理系统的实证”（The Empirical Case for Two Systems of Reasoning）一文。在此文中，他利用自己在人工智能方面的专业知识，提出了稍有不同的双重过程理论（或据其所称，“双重系统”理论）。[注]1999年，加拿大心理学家基思·斯坦诺维奇（Keith Stanovich）在其著作《谁为理性之人？》（*Who Is Rational*？）中，利用自己对推理中个体差异的专业知识，提出了另一个双重过程理论。[注]丹尼尔·卡尼曼在2002年诺贝尔经济学奖的获奖感言中，公开宣扬自己的双重过程理论版本，该理论的许多方面在他与阿莫斯·特沃斯基的早期研究工作都预测过。[注]还有许多人也对该研究做出了贡献，有些人提出了自己对双重过程理论版本的看法，其他人则给出了批评意见。

在对双重过程理论版本的大部分叙述中，可以找到的一个典型策略，便是用表格陈列出对比鲜明的特点。下面是通常会在这种表格中发现的一些对比。

类型 1 过程	类型 2 过程
快的	慢的
不费力的	费力的
平行的	连续的
无意识的	有意识的
自动的	受控的
关联的	基于准则的
有语境的	脱离语境的
启发式的	分析的
直觉的	深思熟虑的
非语言的	与语言相关的
独立于基本智力	与基本智力相关
独立于工作记忆	涉及工作记忆
与非人类动物共有	人类独有

这些对比的主旨很明确：一个是通常与动物本能和人类直觉相关的特点，另一个则是与高阶的有意识的心理活动相关的特点。换句话说，像通常理解的那样，就是与“思考”相关的特点。乍一看，这种差异与理

解一般的人类心理学及特别的推论高度相关：是的，我们人类既拥有无意识的直觉，又可以进行有意识的推理。因此，双重过程理论似乎是受欢迎的重要的发展，而且如果有的话，还是期待已久的发展。这让人如何拒绝这种理论呢？

好吧，人们可能会拒绝这一提法的含糊之处。上述两例很好地分隔开来的特点，在现实生活中会稍稍交织在一起吗？举例来说，我们都会无意识地进行一些简单的算术推论（类型1特点），但这些推论都是基于规则的（类型2特点）。那么，这些简单的数学推论到底是类型1过程还是类型2过程呢？此外，在这种表格中，许多对比——举例来说，在有意识与无意识的进程之间——可能只涉及程度上的差异，而非种类上的对立。类似这种例子都削弱了两种过程类型明显对立的观念。

总而言之，这整个体系的解释性意义何在呢？人类推论中远不止这两种机制。我们认为，解释人类推论的关键在于恰当地鉴别所有这类机制的共同属性、各个机制独有的特点及机制间的衔接关系，而非在薄弱的理论基础上将之归为两个松散的类别。

尽管如此，双重过程理论通过解释推理为何如此频繁地失效，似乎暂时还能帮我们解决所谓的理性之谜。该理论认为，真正的推理（类型2过程）的确是“合逻辑的”，但是就认知资源而言，代价相当大。如果人们的判断并不具备系统的理性，那是因为其普遍基于成本更低的类型1过程。类型1过程是启发式的捷径，在大多数常见情况下，的确会促成正确的判断。然而，在非标准化的情况下，会导致有偏见的错误答案产生。运用类型1过程仍然是有道理的：缺乏高可信度是为追求日常速度及简化推论而付出的合理代价。此外，类型2推理过程仍然能够对类型1直觉过程的结果进行复查。才智上的机敏——如果你喜欢的话，也可以叫智慧——随时准备在需要时被类型2推理过程取而代之。理性之谜解决了吗？并没有。

双重过程理论越发展，就越会鼓励实验研究的发展，而这一简单又

让人满意的图景也就越难以维持了。现在，在埃文斯和斯坦诺维奇看来，把双重过程理论的目的解读为认为类型2过程一定比类型1过程“更好”的观念是个谬论。事实上，他们承认，类型2推理过程本身就能够成为偏见的源头，甚至会在类型1直觉过程产生正确判断时发生错误。我们并未完全回到埃文斯和沃森在20世纪70年代提出的早期双重过程理论，恰恰是因为现在的情况已经丰富得多，但是双重过程理论似乎只是用新的、稍微好点的术语重新提出来罢了。理性之谜仍然存在。

我们不会从任何细节讨论双重过程理论：它是一个太凌乱、变动着且有些模糊的目标。(注)无论怎样，我们都希望给出一些明显更好的东西。与双重过程理论的多样性相比，与我们更相关的是整个双重过程理论的动摇以及推理心理学破碎的方式。几十年来，该领域的核心问题一直都是：人类进行推理时运用的是哪种机制？一些心理学家高喊：“心智逻辑！”另一些则大呼：“心智模型！”有些人仍将此看作核心问题，并利用逻辑学或概率学中的新观点给出了新颖的答复。但这样一来，双重过程理论又有了新疑点。

首先，有人认为，不是一种，而是两种过程类型在起作用。其次，有些双重系统理论家开始认为类型1过程是由许多不同的专门机制共同施行的。最后，就连类型2过程的内部一致性也遭到了质疑。人们越是认为人类推论涉及各种不同层次的机制，“双重过程”及“双重系统理论”的标签也就越难以满足需要。理性与逻辑已经分裂，现在理性本身似乎也已支离破碎。这对于一种研究而言是个不错的结局，对于另一种研究而言也是个良好的开端。

---

1. 要在这个问题上对亚里士多德进行解释并不容易。例如，可参见：Frede 1996。

2. Kant 1998, p.106.

3. 关于实验性推理心理学的不同视角，可参见：Adler & Rips 2008; Evans 2013; Holyoak & Morrison 2012; Johnson-Laird 2006; Kahneman 2011; Manktelow 2012; Manktelow & Chung 2004; Nickerson 2012。

4. 例外情况参见：Evans 2002；Oaksford & Chater 1991。

5. 这些相关问题在哲学家吉尔伯特·哈曼（Gilbert Harman）的著作《视角的改变》（Change in View，1986)中已经得到强调，但仍未对推理心理学产生我们认为应该产生的影响。

6. Khemlani & Johnson-Laird 2012, 2013.

7. F.X.Rocca,“Why Not Women Priests? The Papal Theologian Explains”,Catholic News Service, January 31, 2013，网址：http:// www.catholicnews .com/services/englishnews/2013/why-not-women-priests-the-papal-theologian-explains.cfm。

8. 于2012年11月21日进行的私人交流。

9. 关于这个问题，我们更喜欢Sperber, Cara & Girotto（1995）的解释。

10. Evans 1972; Evans & Lynch 1973.

11. 在解决这个选择问题时，进一步的显著证据证明人们并不进行推理，而只是跟随相关直觉（埃文斯提到的更加丰富的直觉），可参见：Girotto et al. 2001。

12. Evans & Wason 1976, p.485；Wason & Evans 1975.

13. Evans & Over 1996.

14. Sloman 1996.

15. Stanovich 1999.

16. Kahneman 2003.

17. 更多对于双重系统理论的强烈反对意见，可参见：Gigerenzer & Regier 1996; Keren & Schul 2009; Kruglanski等人，2006; Osman 2004.

第二部分
理解推论

象鼻是鼻子的一种。无论象鼻给人的印象多么深刻，将之看作鼻子的典型是说不通的。同样，理性也是推论机制的一种，它并非最好，也不是其他所有机制的典型。为了理解理性，我们必须先理解一般的推论，理解其多样性及影响力。在第三到六章，我们展示了推论在人类及其他动物中的普遍性，也着眼于其基本机制与过程。正如象鼻可以做到其他鼻子无法做到的事情一样，我们会阐明，高阶的人类推论机制（包括理性在内）如何碰巧拥有无可匹敌的广阔影响。

第三章
从无意识推论到直觉

笛卡儿曾坚持认为，动物并不思考。18世纪的苏格兰哲学家大卫·休谟对此持反对意见。他认为，动物思考并能像人类一样进行推论是很明显的。他写道："因此，动物在这些推论中并没有受到推理的指导；孩子们也没有，其平常的行动与决定也没有体现人类的共性；哲学家也没有，他们在生活中也基本与普通人一样。"他还补充道："自然界一定提供了其他一些具备更充分、更一般性用途的法则；我们不相信在生活中不确定的推理及论证过程可以产生极好的推论，从起因推断出影响亦是如此。"注

不仅笛卡儿，大多数哲学家也认为推理及论证是产生更大确定性的途径，还是进行推论的唯一方式。休谟却不为所动，他认为推理太不可靠了，自然界一定给出了其他一些推论方式。但是，若真有比推理更好的"具备更充分、更一般性用途"的方式，那又为何要如此麻烦地去推理呢？

"推理"与"推论"两个词经常被看作同义词。休谟想表达的是，推理只是进行推论的一种方式而已，还不是那种可靠的方式。我们认同这个观点。

为何以及怎样

什么是推理？每个人对此至少有一定的了解，毕竟这是我们都在做

且有意识在做的事情。但是，对某种机制有一定了解是一回事，要真正理解这种机制到底是做什么以及怎样做的却是另一回事。你知道如何吞咽，可是你知道吞咽过程是如何进行的吗？要理解推理（而非仅对其加以利用），人们先需要学会有效区分推理与其他心理过程。而这就是困难的来源。

在哲学及心理学文献中，人们通常用两种方式来定义理性，即其目的或过程。

但可惜的是，这两种释义所指的并不是同一种现象：对目的的标准表述指的是一般性推论；对过程的标准表述指的是正当理由。

我们为何推理？据说推理的目的是通过利用已获取的信息来得出新结论，而非单单凭借观察或他人给出的证据。

我们如何推理？推理的过程表现在为新结论寻找支撑的理由。

比如说，你会犹豫晚上在家看小说还是去电影院。你或许会在某一时刻发现自己不自觉地看起了小说，而你当时并没有仔细考虑该做什么。你或许会想：“今晚没什么好看的电影，天气也不好，要是去了，我可能就得冒雨走回来。另外，这本小说是特默克（Tomoko）给我的，而且看起来真的不错……”如果这就是你的思路，那么你待在家的决定就涉及了推理（暂且不论你是真的运用理性做的决定，还是只是运用理性来对决定进行合理化）。

我们在推理时，结论不是不证自明地在我们的脑海中突然出现的，我们是想到了支撑的理由才会得出该结论。或者，如果我们已经接受了一个给定结论，我们仍然可能会用到推理，因为我们要找到佐证这一结论并能说服他人的理由。这完全是对推理过程的基本概述。不像大多数理论，双重过程理论甚至都没有提到逻辑在确定理由时起到的作用（很多人都认为这是有必要的）。因为它避开了逻辑性框架，说的不是前

提，更广泛地讲，是理由。尽管如此，它已经提出了一个简单的问题：这一过程（关注理由）是从我们已经获取的信息中提取新信息的唯一方式吗？休谟有力地回答道：当然不是！毕竟动物也会预测未来。它们的生命就取决于这些预测是否完全正确。由于未来是无法感知的，动物便通过推论来进行预测。不过，动物是不太可能为其预测找理由的。

我们人类真的就与其他动物如此不同吗？就像它们一样，我们也无法感知未来，只能推测未来；像它们一样，我们的许多日常行为都基于自身欠考虑的预测。举例来说，打网球时，为了把球打回去，你需要适时调整自己的站位；你不需要深思熟虑，一瞬间就能推断出最好的站位。又比如你在与父亲通电话时，仅凭其声调你就能推断出他现在心情不好。你忍住没告诉他自己参加不了下次家庭聚会。你不用细想就能知道父亲的心情，可以想象，你一旦告诉他自己去不了他可能会变得情绪激动，所以你便没有提到这件事：在这种情况下所有这一切似乎立即就变得明朗起来，所以你采取了相应的行动。正如其他动物一样，人类无意识地、不加思考地就能进行预测，得出各种推论。

效仿休谟，我们会用术语“推论”来指从已知信息提取新信息的行为，而无论其过程如何。[注]我们会保留术语“推理”来指通过诉诸理由达到这一目的的特定过程。我们认为，一方面，人类清醒时没有哪一刻不是在进行推论的；另一方面，他们几小时甚至几天不推理都行。

既然诉诸理由已经被认为（更确切地说，被误认为）是从已知信息提取新信息的唯一方式，那么就没有必要或动机来区分“推论”和“推理”了。因此，这两个术语长久以来一直被当作同义词也就不足为奇了。

即便在今天，不认同休谟看法的许多哲学家还是想要沿用“推论”的狭义概念，将之与传统意义上的推理或多或少当作一回事。这就使他们用“推论”与“推理”两个词来表述同一事物，却没有词语来表述其理由并

无因果作用的推论的众多形式。除了保守之外，这种毫无帮助的术语学策略似乎也没有明确的动机。

注意，在某些情况下把“推论”与“推理”看作同义词无伤大雅，甚至还可以达到某种目的。譬如，心理学家经常使用“推理”这个词来形容动物及婴幼儿进行的非凡推论。有些哲学家坚持缩小“推论”的意义范围，使之与“推理”的意义相同，与这些哲学家不同，比较与发展心理学家所做的是扩大“推理”的使用范围，使之涵盖各种推论。使用“推理”这个听起来更优越的词可以表达对生物（动物及婴幼儿）智力应有的尊重，毕竟他们常常被认为并无发言权。

可是，这里我们想要探讨的是推理在其他推论形式中十分特殊的位置。为此，我们最好对二者进行严格区分。当人们不仅在推理中找寻推论，还在可能出现的任何地方找寻时，便能在任何地方发现其踪迹——低等动物、婴儿及成人中，即便他们根本不是在进行推理。

沙漠行军蚁

达尔文看了休谟对动物推论的评论之后，在自己的日记中评论道，人们应当“把理性起源看作是逐渐发展起来的”，最初的提示就是应该从进化的视角看待理性。[注]人类理性的演变用了100多年才成为一个严肃的研究课题，特别是在进化心理学这个新领域。不过，在此之前，达尔文的观点就激发了动物心理学研究。在探究是什么让动物种类得以区分的过程中，“比较心理学”（这门新学科就是这么命名的）突出了所有认知系统的共同之处——它们都进行推论。

以沙漠行军蚁为例。[注]苏黎世大学生物学家鲁迪格·维纳（Rüdiger Wehner）花了30多年时间致力于该项研究。他解释道：

撒哈拉沙漠中的盐滩——平坦、炙热而干燥的广阔地带——只有极少数动物在那里栖息。沙漠行军蚁作为灵巧且活泼的蚁类，无疑是这些物种中最令人瞩目的。它在沙漠表面猛冲、跳跃、攀爬，扫荡遍及各处的食物残渣——主要是死于恶劣的沙漠环境下的其他昆虫的尸体。沙漠行军蚁会离开地下栖息地，在方圆200米开外的沙地以迂回的方式觅食。一旦找到一丁点儿食物，沙漠行军蚁就会牢牢抓住，“打道回府”，几乎径直跑回觅食之旅的起点——通向地下巢穴的微孔。㊟见图3.1。

沙漠行军蚁是如何在归程中准确定位，并在到达巢穴附近时停止高速前进的呢？通过众多周密的实验与仔细的观察，维纳与其合作者指出，这些蚂蚁的“导航工具箱”（navigational toolkit）中有能够估测方向偏离程度的“天文罗盘”（celestial compass），还有能够测算两次改变方向后的距离的“里程计”（odometer），毋庸置疑，这两个任务都远非单纯记录感官信息而已。

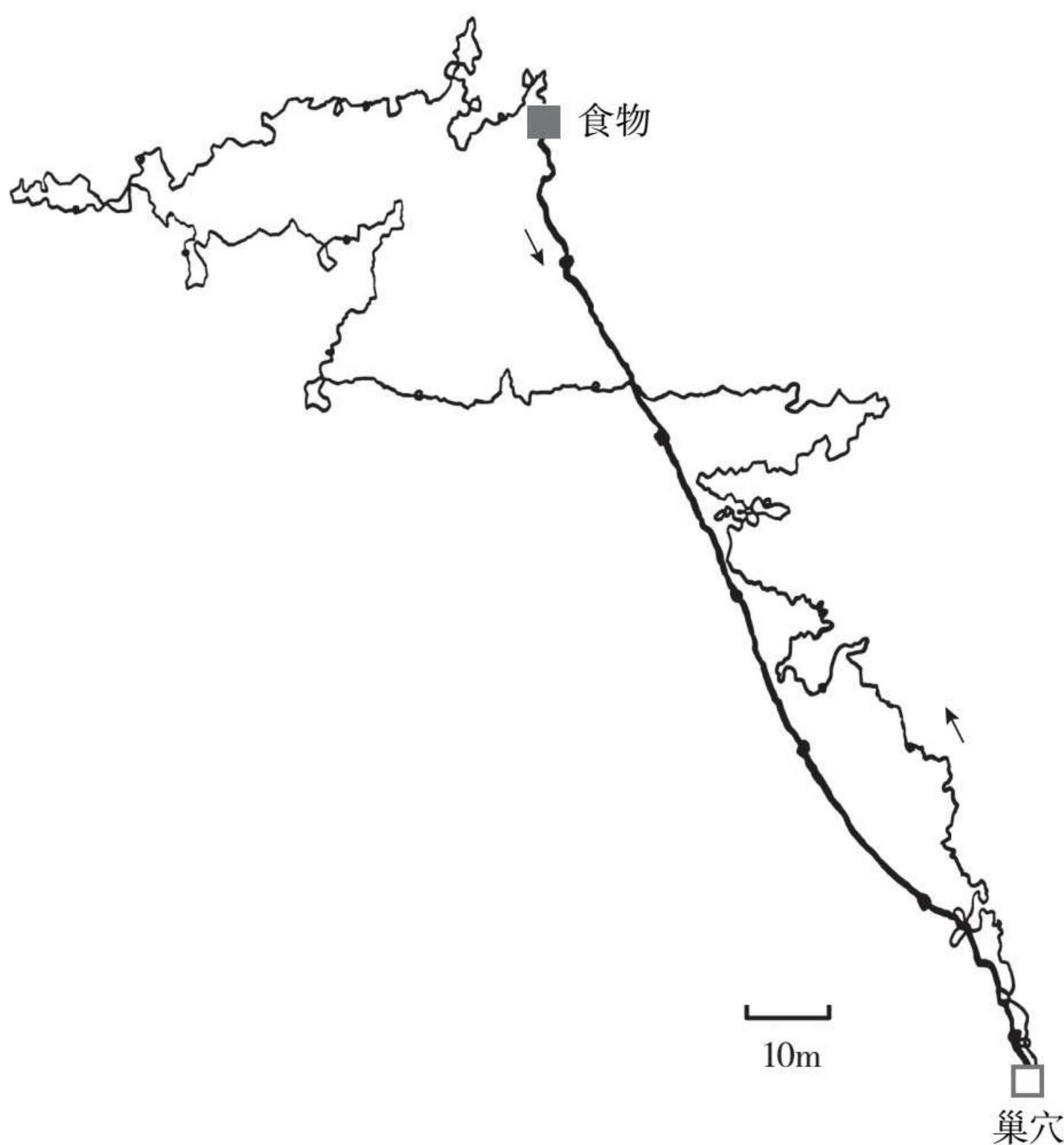


图3.1 沙漠行军蚁从巢穴出发觅食及返巢的路线

“天文罗盘”利用沙漠行军蚁对太阳光偏振的敏感度来确定轴线，并记录沙漠行军蚁相对于此轴线每次向外行进的分割角度。“里程计”则基于走完路程所需要的步数来推算两次改变方向后的距离。之后沙漠行军蚁会在大脑中将“天文罗盘”与“里程计”推算出的所有信息整合到一起，进而推算回到巢穴的距离与方向。这个“路径整合”过程与海员、飞行员在无地标情况下为估算航船或飞机方位所运用的航位推测技术类似［如

今则可与GPS（全球定位系统）的定位输入相媲美］。虽然人类的航位推测法是复杂的智力任务，需要借助测量仪器才能完成，但是沙漠行军蚁的路径整合是在微小的大脑内通过这些自动且无意识的计算来完成的。

鉴于沙漠行军蚁的特质，其觅食的例子可以很好地阐明认知系统的以下三个基本特性。

第一，认知作为工具，让有行动能力的生物体可以恰当应对环境中的风险、挑战。无行动能力的植物并无认知，只有有行动能力的动物才有认知。无行动能力的认知是无用的，没有认知的行动能力则是致命的。沙漠行军蚁，特别是爬出巢穴的沙漠行军蚁，若无认知技能，很快就会在灼灼烈日下毙命。

第二，认知所需要的远不止靠感官就能获知的信息。所有感觉器官通过信息形式获得的，无论是在蚂蚁中还是在人类中，都是成千上万神经末梢的能量变化。要整合这一信息，确定环境中促成这些感官刺激的事件并对这些事件做出适当的回应，认知在很大程度上一定在对事物面貌、期望值及解决方案进行推论。觅食的沙漠行军蚁每分每秒都在推论。

第三，推论可能是由专门的机制来执行的，其中每个机制都尤其擅长处理一种十分具体的任务：依据步数推断距离，依据太阳方位推断方向的角度变化，依据距离和方向推断返巢的最佳路径等。

船上的托勒密

人类有专门的机制且每种机制只处理一种认知任务吗？当然有了，最先就是知觉。众所周知，知觉是由专门机制（视觉、听觉、嗅觉）来实现的。可是，这些机制在推论时没那么明显。我们平常的知觉只是对

事物面貌的记录与对事实的“吸纳”，而非对心理表征的建构。即便存在有关知觉的推论，通常也都是无意识的。

在认知心理学教科书中，通常认为19世纪德国科学家赫尔曼·冯·赫尔姆霍茨（Herman von Helmholtz）是发现知觉中无意识推论的学者。然而，这一发现其实有更深的渊源，公元2世纪希腊-罗马（Greco-Roman）的科学家托勒密(Ptolemy)才是谈及视觉中无意识推论作用的第一人，同时，他还是将错觉当作心智运作关键性证据的第一人。[注]

你可能有过这种体验：自己乘坐的火车其实已经开动，可是自己并未察觉到，甚至一瞬间认为开动的并非是自己乘坐的这列火车，而是对面站台的那一列火车。托勒密生活的年代还没有火车，于是他描述了发生在船上的类似错觉。他写道，假使“我们乘坐的船沿着海岸在行进……可我们并未感觉到它在移动，由此我们断定是海岸线上的树木及地貌在移动。错觉就源自……我们推断可见物在移动”（加上了强调）。

对许多错觉的最佳解释是，它们是源自我们无意识做出的、用来理解自身感觉的推论。这些推论在大多数情况下都是正确的：若我们的位置相对于周围物体发生了改变，那么不是我们自己在移动，就是周围物体在移动。如果我们并未觉得自己在移动，那么推断周围物体在移动也是合情合理的（即便并非万无一失）。

在现代心理学中，错觉为知觉研究提供了关键证据。我们在第一章中看到了阿德尔森棋盘错觉提供的显著例证。罗杰·谢巴德（Roger Shepard）打造了另一个惊人的视错觉事例（见图3.2）。[注]请看这张描述隧道中一只怪物追逐另一只怪物的图片。追逐者看起来比被追者更大，对吗？其实，你可以测量出二者恰好大小一样。所以，投射到视网膜上的图像也应是大小一样的。可为什么一只怪物看起来要比另一只怪物大呢？因为我们看到的不仅是视网膜刺激感知到的物体大小；为推断

物体大小，我们无意识地使用了与估测物体距离相关的背景证据。在这张图片中，两只怪物的肖像与你的双眼是等距离的。尽管如此，因为我们自然而然地把这张图片当作对三维立体场景的表征，鉴于追逐者处于被追者后方，所以它离我们更远，由此我们便认为追逐者更大。凑巧的是，我们在看这张特别的图片时，运用的信息具有误导性，持有的假设也是错误的，我们产生了视错觉。可是，“这个错觉出人意料”的事实告诉我们，通常情况下我们是如何踏实放心地依赖这种无意识推论来获得真正感知的。大多数情况下，此举无可非议。

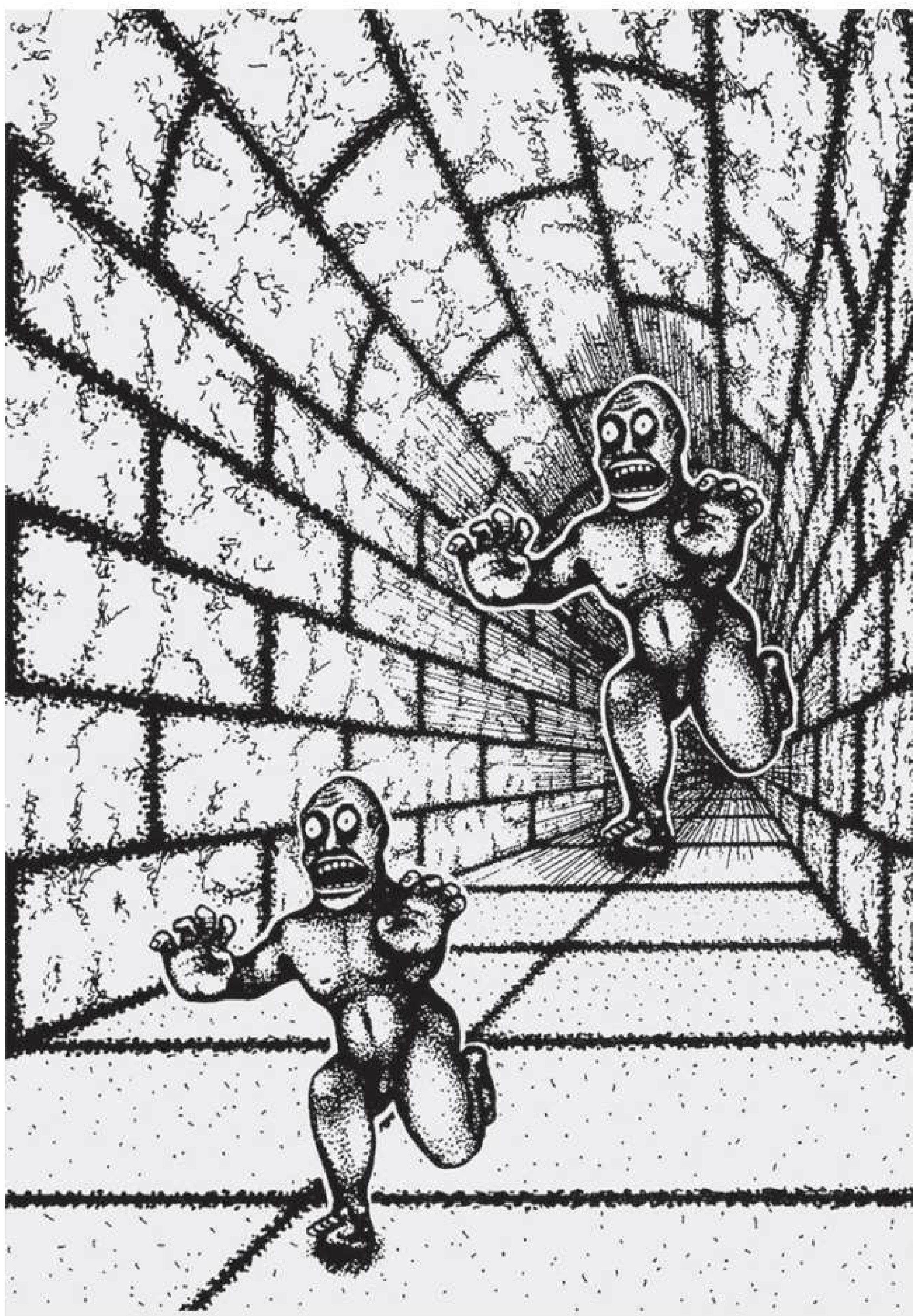


图3.2 隧道中的怪物

知觉所涉及的推论过程通常都十分迅速——其持续时间以毫秒计算——所以我们完全没有意识到。知觉似乎与现实有某种即时、直接的联系。而推论既需要用到时间或空间上相距甚远之物的信息，又免不了犯错的风险，似乎与这种直接联系不相称。好吧，知觉中也有犯错风险；错误的知觉与错觉确实也会发生；我们的知觉是先前经验的产物，有时也会被先前经验误导。所有这些都表明，与消化、呼吸等直觉相比，感知等“直觉”不应被赋予更大的权力。我们对自身心理及其他生物过程存在有意识的直觉（或者说“内省”），但这些直觉没有提供可靠的表述，更不必说对这些过程的解释了。如同所有直觉，这些直觉本身就需要得到解释。

尽管如此，你也许会反对，把知觉中完全无意识、超高速的过程与推理中有意识且缓慢的过程（有时慢得令人痛苦）都放到“推论”这个范畴下又有什么用呢？这不是很勉强吗？就像坚持认为应该把跳跃和飞翔当作同一个“空中移动”过程中的两个例子来学习一样。事实上，这个反对意见的作用并不大。跳跃与飞翔是截然不同的，而知觉中的无意识推论与推理中的有意识推论则处于连续统一体的两端。在二者之间，有许多分工明确的推论过程。有些推论过程快些，有些则慢些；人们在这些推论过程中或多或少地意识到了一些思绪正在酝酿。

虽然所有推论机制都处于一个连续统一体中，但并不是说这些推论机制就是一样的。这一事实表明，虽然这些推论机制共享推论功能，但彼此间的差异可能很大。推理又如何呢？推理只是这许多推论机制中的一种罢了。

我们并未意识到的推论

当你看到这张描绘隧道中两只怪物的图片时，不只注意到该场景的

视觉特性（正如我们所言，这已经涉及了某些推论），你还对自己的所见进行解读。举例来说，你认为这两只怪物正在奔跑（而不是一只脚静止站着）。你认为一只怪物正在追另一只怪物（而不是试图模仿其运动）。你认为追逐的怪物一定有敌意，而被追的怪物一定很害怕。即便两只怪物的脸完全一样，你对二者的解读也不相同。

知觉在解读我们感知到的到底是什么时，可能会有一定的自由度。虽然我们只看出一只怪物比另一只大，也不会轻易被人说服，承认自己弄错了，我们更愿意相信这两只怪物可能被第三只更大且没有出现在图片中的怪物追逐着，而非我们的第一种解读——一只怪物追逐着另一只怪物。我们自然而然地想到了第一种解读，但我们很容易就能注意到，这是一种解读，不是单纯地陈述事实。

记忆同样需要推论。“储藏在记忆中”这一表达，让人想到记忆就像一个储物间，我们可以把物品安全保管在那里，也可以在需要时从那里取出，但这是一个颇具误导性的表达。英国心理学家弗雷德里克·巴特利特（Frederick Bartlett）在1932年出版了一本现今仍有影响的著作《记得》（*Remembering*），书中他介绍了再生产记忆与再构造记忆之间的经典性差异。[注]如果你需要记住一串乱序的数字，会借助死记硬背，必要时你的确会尽力在脑海中重现这串数字。但这根本不是记忆在大部分时间里的典型工作状态。巴特利特写道，我们应当摒弃这样一种观念，即认为“记忆主要是（字面上的）重复或重现。世界无时无刻不在变化，字面上的回忆变得微不足道”。[注]那么，我们是如何记忆的呢？

正如知觉机制通常能通过知觉错觉得到最佳展现，正常的记忆机制通常会在受到诱导进行错误回忆时得到展现。举例来说，布伦特·思特里克兰德（Brent Strickland）与弗兰克·凯尔（Frank Keil）向人们展示了踢球或掷球的一些短片。[注]在半数短片中，触球（或释放球）瞬间被删去了。参与者看完每部短片后，立马就要遵照指示去看一组静止的照片，并要指出这些照片是否在短片中出现过。当短片中整个事件序列，

特别是球的运动暗指触球瞬间一定发生过时，绝大多数参与者都“回忆”起自己看到过这个触球瞬间，即便事实上它并未在短片中出现。确定无疑的是，参与者推论性地重构、“回忆”了必须发生过的事件序列，而非其实际所见。

同样，迈克尔·米勒（Michael Miller）与迈克尔·加扎尼加（Michael Gazzaniga）在实验中给实验参与者展示了细节繁复的彩色图片，描绘的是美国生活的典型场景，如食品杂货店、谷仓场院或海景。[注]起初的图片包含许多典型物件。举例来说，海景中有沙滩球、沙滩毯、沙滩伞及救生员的救生工具。参与者实际看到的图片则经过了修改：两件这类典型物件已经被移除（对于不同组别的参与者，所移除的两件物件有所不同）。米勒与加扎尼加猜测，人们会“记起”自己实际未看过的物件。

参与者看过图片半小时后，会听到实验人员念物件清单，并遵照指示判断所念物件是否在看到的图片中出现过。确实，他们记错了所见之物，他们以为自己看到了沙滩伞及救生工具，而这两个物件已经被移除。参与者回忆起沙滩伞、救生工具的比例就跟他们回忆起有沙滩球和沙滩毯的比例一样高，而后者是真的出现了的。怎么会这样呢？

在上述所有例子中，回想起正确的和错误的记忆都需要用到推论——举例来说，对踢球脚法的推论，因为这可以解释足球随后的运动轨迹，或者是对海景图片中“一定有什么”的推论。通常这种推论都是完全无意识的，而且回忆起来似乎迅速且不费力。不过，实验参与者有时会犹豫一下：图上真的有沙滩伞吗？但这种疑虑很快就消失了。怎么会这样呢？推断出的物件大部分时候都是正确的，但并非一直正确。

在知觉与记忆中，推论一直都在起作用。大多数时候，我们完全没有意识到它的作用。仿佛我们感知到的就是即时呈现在我们面前的东西，我们仿佛是在检索自身记忆，就像它们被存储了一样。不过，我们有时也会意识到自己已对所见之物进行了解读，或是对自身记忆进行了

重构。知觉与回忆失去了某些明显的即时性与清晰度。在这种情况下，我们意识到，自身感知与回忆涉及某些直觉性顿悟。

推论或多或少是有意识的——或者说，或多或少在某一时刻变得有意识，这一点在言语理解上能得到更好的阐释。假设此刻你正坐在一家咖啡厅，无意中听到邻桌一位女士跟坐在她身边的一位男士说“这是水”。要理解这个简单的句子当然没有问题，但你仍然不知道这位女士说的是什么意思。哲学家保罗·格莱斯（Paul Grice）坚决认为，句意与说话者的意图截然不同。注

这位男士可能把手指向自己衬衫上一块湿了的地方，而一旁的女士则安慰他说“这只是水”罢了。她也有可能是用夸张的手法，来抱怨自己的茶水太淡了——“这是水”。也有可能她的语义与即时的情境无关，他们可能在讨论地球面临的最大威胁是什么，她可能是在说“是淡水供应匮乏”，等等。

与你不同的是，与这位女士交谈的人明白她的意思。不过，不是因为他的英语水平更高，而是因为他有相关的语境知识，即他们之前谈过的内容，或他们对彼此的了解，或他们碰巧共有的经历或共同持有的看法。有了这一语境知识，以及该女士所用词语的语义给出的碎片化指示，要推断女士的意思，他的处境更为便利。举例来说，如果男士知道这位女士喜欢喝浓茶，看见她小啜一口后皱了皱眉，那他当然能理解她的意思——她是嫌这茶太过寡淡。

大多数时候，理解所需的推论都在似乎毫不费力、对该过程无意识的情况下完成了。就好像我们只是从对话者的话语中领会了她的意思。然而我们有时却会犹豫。故事中的男士可能并不知道身边的这位女士喜欢喝浓茶，可能会经历片刻的困惑，才能领会到她的意思。然后，他便会意识到，他必须推断出该女士想要表达的意思。理解总是需要用到推论，即便大部分时候我们并未意识到这一点。

直觉

直觉与普通的知觉形成对照，我们认为普通的知觉就是单纯地陈述事物面貌，却没有意识到知觉所包含的推论工作。举例来说，在对隧道中两只怪物的视错觉中，看到一只怪物比另一只怪物更大似乎只是在单纯地陈述事实。另外，将这个场景理解成一只怪物在追赶另一只怪物，会被视为更加脑洞大开的解读。若问你为何觉得一只怪物要比另一只怪物更大，你可能会回答："这是我看到的。"若问你为何觉得一只怪物在追赶另一只怪物，你可能会回答："直觉上显而易见。"

同样，在咖啡厅邻桌的语言交流中，如果这位男士把女士的话"这是水"理解为他衬衫上的那块湿只是因为滴上去一滴水，那么他只是理解了字面意思。如果他进一步认为这位女士是在暗示他：既然这只是水，他就不必担心，那么他对该暗含义的解读似乎就是直觉了。

同样，直觉也与有意识推理的结论形成对照，在有意识推理中，我们知道（或认为自己知道）自己如何以及为何得出这些结论。假设有人告诉你隧道中两只怪物的图像实际上大小一样，你经过测量，证实确实如此，接着你就会接受这个非直觉性结论，因为你有这样做的理由。这是推理性知识，而非单纯的直觉性知识。

咖啡厅里的这位男士也可能会推理："她为何要用这种傲慢的语调来告诉我'这是水'？因为她认为我担心过头了。噢，她是对的——我居然在担心区区一小滴水！一滴水算什么。它干了以后不会留下任何痕迹。我不应该太担心。"当他得出这个自己不应该太担心的结论时，就会去关注支撑该结论的理由。其中便涉及某种推理。

定义直觉的初级方法就是，把直觉当作我们做出的、认为有正当理由的判断（或也能变得颇具直觉性的决定），虽不知为其辩护的理由。人们通常认为，直觉就是"在不知道自己是怎么知道的情况下知道了"。

我们一连串有意识的想法在很大程度上就是“一连串直觉”。直觉在我们的个人经历和一般性的思考、谈论心智的方式（即“大众心理学”）中发挥了主要作用。

大众心理学中一个很普遍的说法就是，我们众多形形色色的直觉是通过一种叫作“直觉力”（intuition，单数形式）的综合能力来传递的。直觉被认为是人们拥有的一种有程度差异的天赋。在这方面，有些人被认为更有天赋，比他人拥有更强烈的直觉。举例来说，人们通常都会认为女性比男性的直觉更准。但直觉真有这种一般官能或机制吗？

知觉不是一种官能的产物，而是好几种不同知觉机制（包括视觉、听觉等）的产物，这显而易见。另外，日常经历并不能告诉我们，在自己各种各样的直觉背后是否有一种单一的一般官能。而且，没有任何科学证据表明，直觉力是一种官能。得到证明的是，我们的直觉是由各种专门化程度不同的推论机制传递出来的。

那么，假设存在许多能传递直觉的推论机制。这些机制有一些共同的基本特性，可以让其既与知觉推论机制相区分，又与推理相区分吗？事实上，通常定义直觉推论时所采用的是其不具备的特性，而非具备的特性。这在丹尼尔·卡尼曼“三种认知系统”的图中表现得尤为清楚，这三种认知系统分别是知觉、直觉和推理（直觉和推理是双重过程理论和双重系统理论里的系统，见图3.3）。注

	知觉	直觉 系统一	推理 系统一
过程	快的 平行的 自动的 不费力的 关联的 学习迟缓 情绪化		慢的 连续的 受控的 费力的 基于准则的 灵活的 中性的
内容	认知 现时刺激 刺激束缚型	概念表征 过去、现在和未来 语言唤起型	

图3.3 丹尼尔·卡尼曼的“三种认知系统”

在这张图中，推理是以其过程特性（即慢的、连续的、受控的等）来定义的，知觉是以其内容特性（即认知、现时刺激、刺激束缚型）来定义的。

而对于直觉的定义则是，运用与知觉相同的过程，产生与推理相同的内容。虽然把所有既不能算作知觉也不能算作推理的推论纳入“直觉”或直觉推论很方便，但被做此解读的类别便显得有点多余，没了自身的积极特性。这会使人对其理论意义产生怀疑。尽管如此，但这并不是问题的最终结果。

若直觉至少在大众心理学中独树一帜，不是因为其是由不同类型的机制创造出来的（大众心理学对此知之甚少或根本不知道），而是因为人们是以一种独特的方式来体验直觉的。当我们产生直觉时，我们把它当作自身心智创造出来的东西，但对其产生过程毫无经验可言。换句话说，即便直觉不是一种基本的机制类型，也很可能是一种独特的“元认

知”类别。

“元认知”，或者说“对认知的认知”，指的是人类评估自己心理状态的能力。㊟比如，当你回忆起把钥匙丢在哪里时，你同样对你“知道钥匙在哪里”有一种或强烈或微弱的感觉。当你从朋友莫莉的面部表情推断出她现在很沮丧时，你也或多或少感觉自己是对的。你自己的认知状态正是元认知评估的对象，元认知评估的表现形式可能为单纯的元认知感觉，在某些情况下，也可能是你对自己思路的清醒思考。

加拿大心理学家瓦莱丽·汤普森（Valerie Thompson）的主张令人信服，她认为直觉有十分独特的元认知特性。㊟我们想要做出更有力的声明，即直觉唯一明显的特征是元认知性。

人们把直觉当作心理状态的一个特殊类型。直觉的内容是有意识的。自相矛盾的是，“我有一个直觉，但我不知道是关于什么的”。另外，人们没有意识到传递直觉的推论过程。事实上，“直觉像是突然出现在意识中”这个事实便是其元认知性特征的一种体现。不过，人们并不认为直觉只是“拿不准”的想法或纯粹的猜测。直觉在产生时也伴随着一种元认知性自信，这种自信或多或少令人信服：人们认为直觉时强时弱。我们对支撑自身直觉的理由知之甚少或一点儿也不知道，但认为这些理由的存在是理所当然的，且理所应当充分地佐证我们的直觉，至少在某种程度上是如此。我们产生直觉时也伴随着一种代理身份或作者归属感。虽然我们并非自身知觉的创造者，却是（大概也能感觉到）自身直觉的创造者，我们甚至还为此感到自豪。

所以，与其把直觉当作“直觉推论”或“直觉”这种特殊的推理过程类型（按照卡尼曼的说法，是过程而非结果）创造出的心理表征，不如把直觉推论当作其产物碰巧就是直觉的推论，这样更说得通。如此看来，“直觉推论”就处于“无意识推论”与“有意识推论”之间。区分这些推论的并非所涉推论机制的特性，而是推论过程及其结论被（不被）元认

知的方式。

拥有直觉的元认知性体验有时是生动而显著的，有时不过是在做判断或决定时难以捉摸的自信心。从完全无意识的推论到有部分意识的推论是一个连续统一体，而且没有办法对这一连续统一体划定明确的界限。但是，直觉与推理结论之间存在界限吗？就推理而言，不只结论，推论过程也是有意识的吗？这还很难说。我们下面将会证明，在推理中，推论过程的晦涩（直觉的典型特征）没有消除，只是替换了而已。因此，直觉的元认知特性就会帮助我们以崭新的视角来看待推理。

---

1. Hume 1999, p.166.

2. 如果说一个有机体拥有关于某些事物现下状态的“信息”，大致就是说这个有机体正处于一种认知状态中，而这种认知状态通常只在上述事物状态存在的情况下才会产生。例如，你通过某种机制获知了外面在下雨这个信息，这种信息获取机制只有在外面真的在下雨时才会产生。这样一种对“信息”的理解与运用从Fred Dretske (1981)、Ruth Millikan (1987, 1993) 和其他有相似想法的作者那里受到启发。我们已经意识到了这一方式引起的复杂问题。尽管并没有必要带着现在的目的去进行深入的探究，但我们仍在第五章讨论“表征”的定义时将此思想有所发展。也可参见：Floridi 2011; Jacob 1997。

3. Darwin 1938-1939, p.101.

4. Steck、Hansson，以及Knaden 2009;Wehner 2003;Wittlinger、Wehner, 以及Wolf 2006。

5. Wehner 1997, p.2.图表可参见：Wehner 2003。

6. 关于对无意识推论等感知的理解，可参见：Hatfield 2002。关于赫尔姆霍茨，可参见：Meulders 2010；关于托勒密，可参见：Smith 1996。

7. Shepard 1990.

8. Bartlett 1932.

9. Bartlett 1932，p.204。

10. Strickland & Keil 2011.

11. Miller & Gazzaniga 1998.

12. Grice 1989.

13. Kahneman 2003.

14. 例如，可参见：Proust 2013; Schwartz 2015。

15. Thompson 2014.

第四章
模块性

休谟的观点是正确的：人类及其他动物一直都在进行推论。但他们在进行推论时用的是同一种机制吗？动物的本能与生俱来，而人类——正如老话所说——只有极少且极有限的一些本能，有高等智力作为补充，且能通过学习来获得知识与技能。但这真的是“二者选其一”的问题吗——要么选择本能，要么选择高等智力和学习？或者说，学习真的可以依赖以不同程度的本能作为基础的、专门的推论机制吗？

本能与专长的博弈

比较心理学家表明，像黄莺、乌鸦或猿猴等动物是通过观察和模仿他人，甚至是通过独自找到解决问题的新方式来获得复杂技能的。发展心理学家则表明，人类进化而来的强烈的性情会在其出生之时就影响认知过程。再举一个例子，人类从出生之日起（如果不是从在子宫内就开始的话），在婴幼儿时期就特别关注人们说话时的声音，并开始学习妈妈讲话。[注]这足以表明人类存在“语言本能”［史蒂芬·平克（Steven Pinker）有句名言，这种“语言本能”远远不止让婴幼儿注意到说话声那么简单］。[注]

为了填平本能与学习之间的鸿沟，个体生态学研究者及鸣禽专家彼得·马勒（Peter Marler）提出，动物也有“学习本能”。[注]根据不同的解读方式，既可以认为该表达自相矛盾——因为学到的不是本能性的，本

能性的不是学到的——也可以认为这是一种原创性方式，阐明了下面这个相当平常的观点：有些动物，特别是人类，具有生物学上的遗传倾向来促使他们学习。不过，马勒本人对这个表述的理解更具体、更有意思。他的意思是，学习本能并不是一种什么都学的不加选择的倾向；学习本能是一种经过进化的倾向，用以习得给定类型的知识，诸如鸟鸣声（对于鸟类而言）或语言（对于人类而言）。学习本能不仅是针对具体的学习目标，它还为本能性学习者提供合适的知觉及推论机制，用以从特定种类的迹象中提取出特定种类的知识。

本能可以看作“自然性专长”（natural expertises）。专长可以看作“习得性本能”。马勒“学习本能”的思想成为洞见起源，是因为该思想认为，在完全进化的本能与完全习得的专长之间存在一个连续统一体，而非需要填平的鸿沟。在这个连续统一体中，认知机制可能占有许多不同的位置。特别是在心理学家研究人类认知机制时，问题不应当是：这是先天性还是后天习得？问题应该是：在每个个体中，有多少种这样的机制？它们是如何由进化而来的能力与倾向促成的（是出生之时便有还是后天发展成熟的）？一些进化而来的学习能力在多大程度上是在针对具体的学习目标呢？反过来说，这些能力又是在多大程度上促进人类学习多种领域的知识的？虽然这些领域的知识各不相同，但恰好能运用同一种“解释模式”(注)来达到最佳的学习效果（正如我们自然地运用心理学范畴的知识来学习并思考个体、团体及组织）？学习本能是如何利用经验来造就成熟的认知机制的？

人脸、规则与书面语言

对于人脸认知机制，一直存在着无休止的辩论。在人类社会生活中，识别他人脸部的能力有重要作用。这种认出人脸的能力只是来源于婴幼儿时期的大量练习吗？这种练习受到很大激励是因为认出他人便会

获益、无法认出他人便会付出代价吗？或者说，也有一种在某种程度上对于处理人脸信息异常高效的进化倾向，这种进化倾向能够驱使婴幼儿将注意力投向人脸，并为其提供一种固有的“人脸模板”及利用感官输入的程序吗？

虽然没有人会否认经验在该问题上对习得成人能力起到的重要作用，但是大量证据似乎表明，存在一种进化倾向，且该倾向的作用至关重要。美国神经心理学家南希·堪维什尔（Nancy Kanwisher）将大脑中一小块区域（在左右脑的下颞叶中）称作“梭型脸部区域”，这一区域在人脸识别中有重要的作用。[注]这一区域若受损，便不能进行人脸识别。人脸识别的有些特性，其他较为复杂的视觉刺激识别并不具备（至少是在同等程度上）。举例来说，如果是识别一张颠倒过来的人脸，效率则要低得多，颠倒方向对识别人脸造成的影响要比其他任何类别的视觉刺激大得多。尽管如此，一些研究者收集的证据及论据表明，并没有专门用于人脸识别的特殊技能。与识别其他东西相比，我们只是对从视觉上识别人脸更有经验罢了。他们认为，人类只是在识别人脸上拥有源于习惯的专长而已。

在我们看来（但我们并非专家），这个例子强有力地说明，特定领域人脸识别机制的个体发展是受一种进化基础引导的，而且这些年来，该观点已经令人信服。即便如此，我们引用这个例子并非为了在这件事上表明立场，而是为了阐明，现今，在充满无限可能性的连续统一体上，从专门的认知性“学习本能”到熟练地运用通用机制来处理基于专家层级经验的特定类型的信息，研究具体推论机制就涉及弄清楚其在连续统一体上的合适站位。上述问题可以有诸多答案的事实本身就暗示着，推论是一种功能，很可能需要通过相当不同的机制来实现。

在与知觉性识别关联更小，而在特性上更为观念化的各种学习及推论类型上，也出现了相似的问题。举例来说，父母通常都非常希望自己还在蹒跚学步的孩子能够学会做各种事情的“正确方式”，希望孩子能够

展示出对刚习得的规范的熟练程度。这就提出了一个难题。当年幼的孩子观察周围人的动作时，他们是如何将示例规则的动作与其他得到社会认可的但未予以奖惩激励的动作区别开来的呢？知觉性线索在这里起到的作用非常有限。而且，跨文化的规则教导的差别很大，更多时候是暗示而非直接言明，永远不会穷尽。那么，年幼的学习者要如何识别自己观察的哪些动作是在示例规则中的呢？

匈牙利心理学家杰格力·西布拉（Gergely Csibra）与加里吉·杰格力（György Gergely）已经证明，婴幼儿倾向于将呈现给自身的“明显示例的”信息，即用有吸引力的、有明确目的的方式呈现的信息，视作其社群中具有普遍相关性的信息。当大人明显示例某些新动作时，婴幼儿很容易就推断出这类动作是在示例“做事方式”。

在一项著名研究中，实验人员给婴幼儿展示直接按压圆顶便可打开的圆顶状台灯。这个按灯的动作由一个大人来演示，但因为用手按太平常，她就用自己的前额来按灯。当把这个不同寻常的动作明显地示例给孩子看时（就像是某种游戏或是仪式一样），婴幼儿很容易就能模仿。然而，如果孩子看到了相同的碰头动作，但不是明显示例的，他们就不模仿了，当轮到孩子来操作这盏灯时，他们就用手按圆顶来打开它。这项研究[注]揭示了孩子有选择性模仿动作的倾向，因为动作是明显示例的，所以孩子认为这是操作示例的“合适方式”。

年幼的孩子能够从明显示例中推断规范的动作。等到年龄更大一些，他们便能够运用动作本身的典型特性来推断其规范性特点。心理学家汉尼斯·拉克西（Hannes Rakoczy）、马科·施密德（Marco Schmidt）、迈克尔·托马塞洛（Michael Tomasello）及菲利克斯·华内肯（Felix Warneken）展示了两三岁的孩子是如何在另一个人没有遵循规范时，自然而然地通过试图施行这个规范来演示自己新近习得的这一规范。[注]这一切表明，人类很可能有一种识别并习得社会规范的类本能性倾向。

人脸识别与行为遵守规范是人类社会生活的普遍特点。因此，这两种情况涉及的认知机制在很大程度上是生物进化的结果，这一说法看似颇为有理。另外，自然选择要进化出诸如冲浪、计算机编程这类新潮实践所需要的专门能力太缓慢了。即便是阅读、下象棋这类更古老的文化性技能，也仍然比为了帮助个体获取知识而进化出来的学习本能要晚得多。

直到最近，阅读都只是少部分专员（学者和抄写员）的技能。即便现在阅读的传播范围更为广泛，但从认知的角度来看，它仍然是一项专长，即通过在有组织的教导及指导下的集中练习来习得的技能。这种专长与只需要一丁点儿或根本不需要学习的普通本能形成鲜明对照。然而，针对阅读的神经心理学研究表明，其神经性基础与像人脸识别这种更为“本能性”的能力颇为类似。

法国神经科学家斯坦尼斯拉斯·德汉纳（Stanislas Dehaene）与劳伦特·科恩（Laurent Cohen）已经证实，个体在阅读中对书面文字的识别用到了他们称为“视觉词形区”的一小块精确脑区，这一脑区在左脑梭型人脸区域的旁边，能够通过个体独立习得的知识，来识别稿本上的字母与词汇是大写字母还是小写字母，是手写还是印刷字体。也有证据表明，盲人用手指阅读盲文时也需要用到这块区域。显然，这种大脑机制提取的信息要比从视觉或触觉刺激中推断出的信息抽象得多。

不同个人、社会及书写系统都在使用的这一小块脑区，竟被用来阅读，怎么会这样呢？德汉纳与科恩的假说认为，阅读者大脑中专用区域的发展是“神经元循环”过程的产物。

> 一方面，读来的东西会“侵入”大脑皮层的特定区域，这些区域拥有合适的接受域，能够识别被当作字母使用的形态各异的小图形，并通过合适的连接方式向颞叶语言区域传送这一信息。另一方面，书写系统的文化形式一定是依照大脑的可学习性限制进化而来

的，进而汇聚到能让这些特定视觉区域进行最理想化学习的一小组符号形状中。(注)

正如这一假说表明的那样，视觉词形区位于专用于知觉性识别具体输入类型（包括人脸）的数个机制所处地带。此外，左脑靠近一旦看到书面文字就要进行解读的语言区域且与其联系紧密。

人脸识别与阅读的例子共同说明，进化性认知技能与专业性认知技能都利用了十分具体的脑区，若这些脑区受到损伤，那么这些技能也会受到损害。在一些重要方面，进化性技能与专业性技能有相似的运作方式：其众多运作都是迅速、准自动化、为具体任务量身定制的。进化性机制与专业性机制在其专门领域中都非常高效。若呈现在这些机制中的输入不属于其领域范围，不过是些刺激罢了（就像云的形状或微笑符可能会刺激人脸识别），那么这些机制的表现可能会比较糟糕，或导致“认知性错觉”。

模块

在本能-专长这个连续统一体上的所有机制，在生物学（或工程学）中可能通常称为模块（modules）：它们是有历史沿革、功能及适于这一功能的步骤的自主性机制。（这些机制应当被看作各自贡献不同、都属于更大系统的组成部分。）另外，若不识别模块系统的各模块组成部分及其共同运作的方式，模块系统的能力就不能得到很好的解释。

模块的生物学概念相当广泛。可以是或大或小的模块，可以是子模块或子模块的子模块（如视觉系统的各个组成部分），还可以是本身就模块化的模块群［就像把神经系统子机制与人手“力量型抓握”（power grip）和“准确型抓握”（precision grip）特点的结合］。整个人脑就是一

个生物学模块，因此也是一个单独的神经元。生物学模块可能是同一物种的所有成员都具有的一个解剖学特点，如象的鼻子或牛的反刍。生物学模块也可能是解剖学特性或行为倾向，只在某些特定的情境下亮明自身，如皮肤反复摩擦产生的老茧，或称为兽群迁徙的集体性行为。

为何许多有机体在很大程度上属于模块系统，这背后有深层次的原因，那就是可能有不同进化与发展轨迹的、相对自主化的机制间彼此关联。单个模块都相对僵化死板，不同模块之间相互关联则能够让复杂有机体拥有适应性的灵活度。一个未模块化的有机体能够有类似的灵活度吗？尚不十分明确。而且，模块系统更有可能克服局部的功能故障，或随着环境的变化进行调整。最重要的是，模块系统有更强大的——有人认为是独一无二的——进化能力。㊟

在心理学中，人们一直以来都认为心智是拥有完整记忆的统一性一般智力，并通过各种感觉及运动器官与世界联系。如今，来自神经科学、发展心理学及进化心理学的证据及论据都表明，心智是很多各式各样的自主化机制的联合体。识别并描述这些机制已经成为认知科学的中心任务。

然而，在心智哲学及心理学中，有关模块及模块性的说法由于一些历史原因存有争议。有人认为，心智可能包括几个自主化模块，这一看法在哲学家杰瑞·福多尔（Jerry Fodor）于1983年出版的开创性著作《心智的模块性》（*The Modularity of Mind*）㊟中得到了广为人知的捍卫。福多尔认为，心智的输入系统（知觉及语言）是模块化的，而其中心过程，特别是推理，则不是模块化的。他对模块严格的但回想起来有些武断的定义，及其对有限的心智模块性的看法，是灵感源泉，也是无休止论战的始作俑者。为了避免这些论战，许多心理学家转而谈论“机制”，或像斯坦尼斯拉斯·德汉纳一样谈论“处理器”。他们放弃谈论“模块”所付出的术语学代价就是，也由此放弃使用“模块”（modular）、“模块化”（modularization）及“模块性”（modularity）。其实，在问及某些问

题时这些概念很有用，如一个给定机制的模块化程度、一项直接教授的技能在什么情况下可能会模块化、进化中模块性的作用是什么。

这里，我们只是将认知模块简单定义为有认知功能的生物学模块，从而完全避免了福多尔怪异的模块性概念引发的有点陈腐的辩论。对于生物学模块与认知功能这两个概念的解读，即便不完美，至少也足以让我们将其整合在一起并进行深入研究。与对生物学模块性概念通常的曲解相反，这并不是说认知模块就一定是“天生的”。我们看到，阅读是认知模块在脑组织中的完美诠释，其生物学进化特性让自身也适合新颖的文化性进化功能。

对于模块性另一个普遍的曲解就是，认为模块化的心智一定是十分僵化的。把单个模块相对的僵硬性与作为整体的认知系统的灵活性（或者是作为整体的大脑的可塑性）进行比较毫无道理。在生物学中，灵活性与可塑性通常都从模块化组织的角度进行解释。我们认为在心理学中也应该是一样的。

我们承认，鉴于该领域的研究进展迅速及余下的研究工作量，我们对人类心智组织当下的理解很可能在不久的将来会得到很大提升，发生很大改变。“心智是各个模块的联合体”这一观念若能得到合适解读，便能作为具有挑战性的可行假设，从而有助于更好地进行解读。注要有所进展，十分关键的是，不仅要更了解推论模块的作用，还要更了解其发挥作用的方式。若是各个模块都相当拙劣，那么要如何组合才能成就更为出色的较大模块呢？每一个人所具备的自信其实是种超然出众的杰出能力，那么又要如何组合才能使最终的较大模块都让人类具备这种杰出能力呢？

---

1. Vouloumanos & Werker 2007.
2. Pinker 1994.
3. Marler 1991.

4. 为了使用弗兰克·凯尔的精确表达(Keil 1992)。

5. Kanwisher, McDermott & Chun 1997.

6. Csibra & Gergely 2009;Gergely,Bekkering & Király 2002;也可参见：Nielsen & Tomaselli 2010。这些研究都以梅尔特周夫（Meltzoff）在1988年的实验为基础。

7. Rakoczy, Warneken & Tomasello 2008; Schmidt, Rakoczy & Tomasello 2011;Schmidt & Tomasello 2012.

8. Dehaene & Cohen 2011.

9. Schlosser & Wagner 2004.

10. Fodor 1983.

11. 例如，克拉克·巴瑞特、彼得·卡鲁瑟斯、勒达·科斯米德思、约翰·图比、罗博·库尔兹班（Rob Kurzban）、史蒂芬·平克和丹·斯珀伯等人的研究中有很好的阐释。更深层次的讨论可参见：Barrett 2015。

第五章
认知机会主义

军队作为一个整体行进时，会尽可能忽视不规则地形，或是将其视作需要克服的困难。相反，自主游击队则会将这样的地形特点视作机会，并且在可能的时候加以利用。驾驶机动船需要考虑风向对航线的影响，并进行微调。航海则需要将风与风向的改变当作能够利用的机会。大致的对比很明显：实现类似的目标可能有时是通过规划行动方案并运用足够的力量来坚持这一方案，有时则通过利用沿途的机会，用更简便的方式继续行进。

经典的推论观点认为，有一个强大的逻辑引擎，无论手边的任务多么古怪，都能够驱动心智很有原则地继续行进在一条笔直的道路上。我们则认为，推论及认知更普通的是通过相对自主化的模块的联合而实现的，这种模块在物种中得到进化、在个体中得到发展，这样的话，它就能在产生问题时解决问题，在出现机会时利用机会。正如游击战争或航海，认知具有机会性。

没有达尔文“物竞天择”的理论（这是机会主义过程的范例），会产生“心理过程是机会主义的”这个想法吗？事实是，只有当达尔文的理论开始影响心理学之后，这一想法才会出现。

然而，早在达尔文之前，无意识推论的发现就对“心智是统一的、有原则的”这一传统观念发起了挑战。第一个正确认识并应对这一挑战的是阿拉伯科学家伊本·艾尔-海什木（Ibn Al-Haytham，也称“阿尔哈森”，Alhacen），他于公元965年出生于巴士拉。他接过了托勒密800年

前交出的接力棒，继续对视知觉的无意识推论进行研究，并进一步发展了这一研究。[注]

伊本·艾尔-海什木猜想

伊本·艾尔-海什木很好奇，无意识推论是怎么进行的呢？用的是与有意识推论相同的方法吗？乍一看，知觉（打个比方）中涉及的即时的、自动化的推论与有意识推理中涉及的刻意的、费力的缓慢推论鲜有共同之处。伊本·艾尔-海什木发现，正如我们在第四章中主张的，在有意识推论与无意识推论之间存在一个连续统一体。他推测，尽管二者差异明显，但有意识推论与无意识推论用的还是相同的工具，即亚里士多德的三段论。在他生活的时代，没有实际的替代性选择。

如今，对于“推论可能是如何进行的”这个问题有好几种颇为不同的表述。有许多不同的逻辑系统。在心理学中，有好几个关于“心智逻辑”的表述，还有约翰逊-莱尔德与拜恩提出的理论，即所有真正的推论都是通过构造并操作心智模型来实现的。概率论推论模型——特别是那些以18世纪英国牧师、学者托马斯·贝叶斯（Thomas Bayes）的思想为根据的——鼓舞了许多新研究的开展。[注]这可能是因为其中有些方式可以很好地描述某具体类型的推论，可是没有一种方式能充分描述一般性推论。不管怎样，这些方式的大多数支持者倾向于认同伊本·艾尔-海什木的看法，即一定存在一种通用的方式可以指导所有形式的推论。但在“这一正确方式可能是什么”的问题上，他们与伊本·艾尔-海什木意见相左，且其内部也存在意见分歧。

伊本·艾尔-海什木意识到，无论这个方式是什么，假定所有推论都使用同一种通用的方式，则会让人产生深深的困惑。同一种方式，怎么可能时而缓慢且费力，时而又不必有意识地花费时间和心力呢？为什么不一直都用那种迅速的模式呢？他的回答是，所有推论最初必须通过有

意识的费力的推理来进行。其中一些推论进行了一次又一次，最终不再有难度；进行这些推论时可以非常快，快到竟会意识不到。因此，他认为，意识程度并不与不同类型的推论相对应，而只与不同的困难程度相对应，最常规的推论最简单且完全是无意识的。伊本·艾尔-海什木坚持认为，从对哲学问题的复杂推理（极少发生）到感知相对大小时的自动化推论（一直都在发生），所有推论都是以同一种方式进行的。

有人认为，起初所有推论都是有意识的，但由于思维惯性，其中一些推论变得无意识。这一想法很有创意，但这是真的吗？很可能并不是真的，因为这蕴含明显错误的预测。如果说迅速且无意识的推论之所以变得迅速且无意识，是因为完全常规化了，那么举例来说，婴幼儿就应该是以缓慢且有意识的方式来推论的。婴幼儿应当在长大后，通过延伸练习才能够达到常规化的自动性。然而，发展心理学家已经证明，婴幼儿自动进行各种普通推论之后好几年，才开始开展刻意的、有意识的推理，这与伊本·艾尔-海什木的解释给我们的预测导向正好相反。

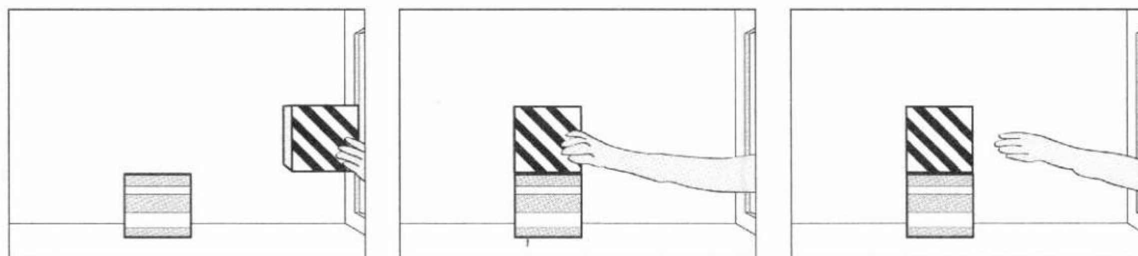
下面举一个例子。心理学家艾米·尼达姆（Amy Needham）与雷尼·柏拉杰恩（Renée Baillargeon）给四个半月的婴幼儿展示了可能事件与不可能事件（见图5.1）[注]。在“可能事件”中，婴幼儿看到有一只手将一个盒子放到了平台上。在“不可能事件”中，婴幼儿则看到那只手在平台外的半空中放开了盒子。这两种情况下，盒子都待在了手放开它的地方。婴幼儿注视着不可能事件中待在半空中而没有掉下来的盒子的时间要更长一些。注视时间上的这一差异充分表明，就像成人一样，婴幼儿也料想盒子会掉下来。

和伊本·艾尔-海什木一样，我们来假设，所有推论都遵循一个逻辑模式，那么就应该能够得出结论，婴幼儿以为无支撑的盒子会掉下来，是因为其遵循了以下这个假言三段论。

前提：1. 一个物体如果没有支撑，就会掉下来。

2. 这个物体没有支撑。

结论：所以会掉下来。



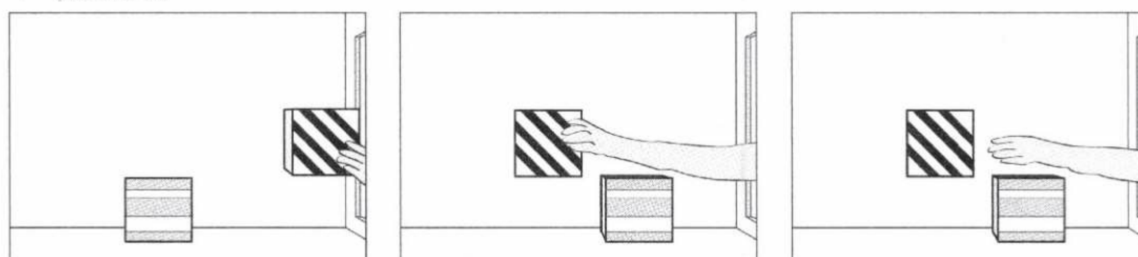


图5.1 四个半月大的婴幼儿看到物理上的不可能事件时很惊讶

此外，婴幼儿也应当是以缓慢、费力且有意识的方式来进行这个推论的（直到年岁增长、经验增加，其推论才能够快到以无意识的方式进行）。然而，这并非心理学家观察到的样子。

这一证据表明，经验确实重要，但与伊本·艾尔-海什木所预测的方式不同。在四个半月大的时候，婴幼儿并不会关注盒子获得的支撑的具体面积比例。即便底部只有15%得到了平台的支撑，他们也希望盒子能够保持稳定。到了六个半月大的时候，他们已经很好地认识到，盒子如果没有得到有效的支撑就会掉下去。注然而，没有证据或论据能够说明，从四个半月大到六个半月大之间的进步，是通过缓慢的、有意识的、费力的推理逐渐变得常规化来实现的。更有可能的是，婴幼儿利用适用于该任务的程序，自动且无意识地提取出统计规律，最终还强化了程序本身。

过去50年来，婴幼儿认知研究成果丰硕，不再有争议的是，婴幼儿能够在推论中考虑到物体的基本属性，而且会渐渐做得更好。值得怀疑的想法是，婴幼儿在预料无支撑的物体会掉下时，利用了有关物体降落的一般性假言前提。婴幼儿真的有这种常识吗？他们在四个半月大到六个半月大之间，通过表征支撑物体所需的、使其不会掉落的面积比例，纠正了这一常识吗？伊本·艾尔-海什木及许多现代作者认为，答案是肯定的：推论必须基于逻辑模式及心理表征这一前提。没有逻辑，就没有推论。

若伊本·艾尔-海什木的想法是正确的，即没有逻辑就不会有推论，那么这一主张难道不应该不仅适用于人类推论，还适用于动物推论吗？哲学家杰瑞·福多尔认为："是达尔文的自然选择确保生物要么了解逻辑元素，要么接受死亡。"注

好吧，还有另一种方式。

表征与程序

所有推论，无论是由蚂蚁、人类还是机器人做出的，都需要用到表征与程序，这一独特之处在人工智能的发展中发挥了重要作用（诸如"数据"与"程序"的对战或"陈述性的"与"程式性的"的对战）。注这与我们对模块化心智进化的理解也有高度的相关性。

先谈谈表征，这是一个让人有诸多困惑的概念。把表征理解为图片或口头陈述很常见。图片与话语是我们在周围环境中较为熟悉的事物，我们会拿出来相互交流。我们同样也把它们当作认知工具来使用。我们运用书面数字进行运算，利用地图来规划旅行，将购物清单作为外部的记忆小工具，等等。

然而，与图片、口头或书面语言不同，我们使用的大多数表征并不存在于周遭环境中，而存在于我们的大脑中；我们运用表征并不是为了与他人沟通，而是为了独立处理信息。同样，在人们看来，心理表征不知怎的一不小心就像是一张图或一段话。我们究竟有没有心理影像？难道我们不是在自身心智中默默地自言自语吗？难道我们不能结合影像与内心语言来思考吗？然而，这种考量难以证明我们是否需要像构造公共表征一样，去构造所有的或者大部分心理表征，又或者，是否有必要构造心理表征。

因此，你可能会问，表征还能是别的什么呢？

我们后面将用到的术语“表征”[注]，是指有形的事物，如大脑中神经元群的活化作用、电子储存媒介中的磁性图或纸上的墨水图案。表征可能在有机体内部，也可能在其周围环境中。使这样一个有形的事物成为表征的并不是其所处的位置，也不是其形状或结构，而是其功能。表征拥有向有机体（或者更广泛一些，任何信息处理设备）提供有关某种事情状态的信息。提供的信息可能是关于事情实际的状态，也可能是关于理想中的状态，即关于事实或目标。

有一个非常简单的例子：在警报系统中使用的运动探测器。质量更佳的运动探测器同时使用两种类型的传感器：一种是微波传感器，能够发射并探测反射回来的微波，这种微波通常会跟随运动着的物体而改变；另一种是红外线传感器，能够探测由有温度的物体发射出的辐射能。联合使用两种类型的传感器能够降低误报风险。激活后，两类传感器会各自发射出电子信号，这种电子信号的功能是告知系统中的下一个设备，激活传感器的事件已经发生了。这一设备对于两类传感器而言功效相同，被称为“与门”（AND-gate）（图5.2）。其功能是推论性的：当通过两个电子输入信息知晓第一个及第二个传感器都被激活时，它会触发一个声音信号，该声音信号会警告人类执行者，受到侵入的可能性已达临界值。

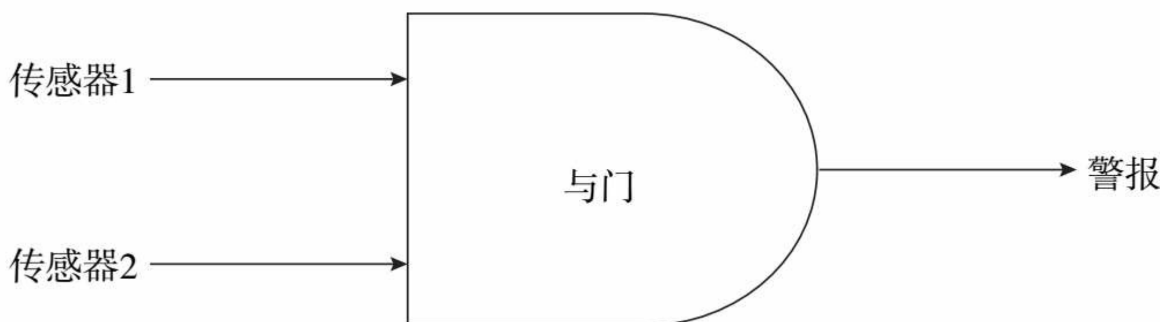


图5.2 在双重科技运动探测器中用到的“与门”

虽然两类传感器的电子信号与声音信号既不像图片也不像陈述，虽然其发挥功能时不需要用到内部结构，但作为“与门”输入信息的两类传感器电子信号，作为“与门”输出信息的声音信号，二者的功能便在于告知电子设备或人类执行者具体的事件类型，因此可以描述为“表征”，这就是我们使用的这个术语的意义。

当然，整个过程都可以用物理学术语来描述，不用谈及信息、功能或表征。然而，要理解人们为何建造、售卖以及购买运动探测器，则需要超越纯粹的物理学表述，从信息和功能方面对设备的功用进行描述，才能够明白。这里，我们以运动探测器为例，用最简单的方式来介绍表征的概念，以此来继续讲述。我们运用“表征”时十分讲求实际：(注)如果不用这个概念（无论是不是有人使用），我们就不知道谈论推论及推理的明智方式。

推论程序适用于表征。它们把表征作为输入，也可以将其清除或予以修正，或者是制造出新的表征（正如运动探测器的“与门”在接收到正确的输入信息时做的工作一样）。正如表征是根据其功能来定义的，推论程序也是如此。使一个程序具有推论性的条件是该程序具有这样一种功能，能够为处理过程提供更多可用信息，或是使已知信息具有更高的可信度。举例来说，推论程序可能在新证据表明其为错误表征时将这一表征清除，可能修改表征以纠正或更新，可能根据其他已知表征制造出新表征，可能增强或降低认知系统对于表征的依赖程度。成功的推论程

序会产生更丰富或更可靠的相关信息。

认知程序应用于心智模块中（以程序在电脑中或是客户端在智能手机中应用的方式）。像反射这种十分简单的程序可能执行的是单一程序，而更为复杂的模块则可能执行并组合了其中好几种程序（而有子模块的模块则可能组合了许多程序）。

心智模块能够执行并使用一个或几个程序（正如电子设备能够执行并使用几个项目，或智能手机能够执行并使用几个客户端一样）。一个模块能够通过与其他模块之间的连接，给其程序提供可以处理的输入信息。为了进行处理，程序可能需要获取某些特别的数据。举例来说，阅读模块所使用的程序需要有关字母形状的信息。这类数据是通过模块为程序所用的，这些模块可能会将这些数据存储在一个专属数据库里，或向其他模块索取。另外，还会与其他相关联的模块共享输出信息。[注]

举例来说，沙漠行军蚁在大脑中将“里程计”与“天文罗盘”的输出信息提供给一个集成模块，而这个集成模块能够估算并更新对蚁穴方向及距离的表征，该表征转而会为运动控制模块所用，用来指导沙漠行军蚁的归巢运动。

好几个模块可能处理相同的输入信息，但将其输送给不同的程序。拥有一个让许多模块并行运作的模块化系统，好处主要在于能同时输出多个结果。毕竟，这是动物需要的一种推论能力，用来监控它们周围复杂的环境，并及时探测到不同的威胁与机会。

哲学与心理学在发展历程中，重心一直都放在有意识推理及其继而使用的明确程序中，用的是学者不慌不忙、全神贯注的方式——想象一个生活闲适、有学术兴趣的绅士，拥有秩序井然的生活，将日常生活的大小琐事与变迁兴废都交给了仆人与女人。从伊本·艾尔-海什木开始，学者将注意力投向发生在视知觉（打个比方）中的无意识推论机制时，他们通常认为，所需要的程序与其自身有意识推理所需程序完全相同或

十分类似，还认为这些程序作用于陈述或类似陈述的表征。然而，这既非必然真理，也非实证经验强有力支撑的假说。伊本·艾尔-海什木在研究中的大胆猜想，已经成为古老的教条。

教条之外

很长时间以来都有这样一个教条，即所有推论，不论是有意识的还是无意识的，都使用同样的亚里士多德式逻辑程序，并将其应用于类似陈述的表征中，这一教条的出现是因为缺乏可替换性选择。推论还能怎么进行呢？直到最近，这都只是个修辞学问题，但现在不是了。该教条已经遭到了形式及实证研究的破坏。

在形式研究方面，自17世纪以来概率论逐渐浮现，以及自19世纪以来现代逻辑学的发展与多样化，亚里士多德式推论模型已被废弃。然而，形式研究发展带来的结果不是质疑“推论必须基于在不同领域间通用的一些小程序”这一观念，而是对这些通用程序是什么展开辩论。

在实证研究方面，对认知、认知进化、跨物种多样性、儿童的认知发展、认知在大脑中的实现的研究，人工智能领域的进步，对认知性过程及脑过程的数学建模，都表明可以采用许多不同的方式来进行推论。推论时可能需要用到各种程序，其中许多程序是专门用来从某个具体的实证领域提取信息，或专门用来执行一种具体类型的推论任务。其中有些程序，除了都具备推论性外，鲜有共同之处。不论存在哪些差异，它们都是从已知信息中找到根据来修正或扩展推论的程序。

某些程序之间很可能存在重要的共性。举例来说，递移推论（“A比B大，B比C大，因此A比C大”）与众多领域都有重大关系，与物理领域、社会领域等都是如此。同样十分合理的是，许多推论程序都在更新未来事件发生的可能性——做出并修改概率论预测，如果你愿意的话

——同时进行微调使自身适应具体领域的规律。注

某些非常专业的推论模块不过是认知反射而已。俄罗斯心理学家巴甫洛夫（Pavlov）在一项著名实验中，让狗在听到铃响时习惯性地流口水，因为每次铃响之后都会有食物出现，所以狗很有经验。研究这种条件性反射对行为主义理论的发展有重要作用，是一种否认或至少忽视心理状态的心理学研究方式。从后行为主义者的认知视角来看，巴甫洛夫实验中狗的条件性反射既是认知性的，也是行为性的。注它促使狗期待食物出现（认知性反应）并分泌唾液（行为性反应）。

狗的脑海中可能没有出现这样的事：没有通用推论程序和亚里士多德式三段论，二者会用作前提的是两个呈现在狗的脑海中的类似陈述的表征，我们可以将之改述为“如果铃响了，食物就会来”和“铃响了”。狗的脑海中并没有“如果……就……”这种可以让假言推论中的演绎推理规则应用于此的大前提。巴甫洛夫让狗形成条件反射的过程是在一个完全专门化的模块里进行的，这一模块利用了铃声—食物这一规律却没有对其进行表征。

每次铃声响起的时候，表征在狗的脑海中就是铃声响了这件事。这一表征告知狗的认知系统，或者更确切地说，其条件性反射模块，铃声正在响起，以此启动程序。这个模块就只有一个认知性影响，就是让狗产生对食物的期待，以及一个行为性影响——让狗分泌唾液。在认知主义者看来，这仍然是一个推论模块：其任务是从观察（铃响了）中得出一个相关结论（食物要来了）。只要铃声—食物规律仍在周围环境中维持着，这一反射推论从认知角度来看就是合理的。

事件若在完全混乱的世界中发生，就无法进行任何相关推论，逻辑会变得毫无意义，概率也不会有任何帮助。让相关推论成为可能的是——不论是物理学家的推论，还是狗的推论——世界上存在可信赖的规律。有些像是物理法则这样的规律很普遍，其他像是巴甫洛夫实验中铃

声—食物这样的规律则十分短暂且有局限性。正是这些规律（普遍规律和局部规律）让我们（包括非人类动物和人类动物）理解自身感官刺激及之前存储的信息。最重要的是，它们让我们能够预测接下来可能会发生的事并采取适当的行动。没有规律就没有推论。没有推论就没有行动。

动物（包括人类）已进化到能够利用周围环境的规律，但它们并没有进化到可以利用所有规律或是普遍规律的地步。企图这么做只是对时间与精力的一种可笑浪费。动物考虑的规律就只是对其成功繁殖很重要的、有时直接但通常间接的规律。

四处移动的动物起初利用的是周围环境中能够帮助或阻碍其运动的物理特性。正在觅食的动物利用与找到食物相关的规律；被捕食者利用捕食者行为中的规律，捕食者则利用被捕食者行为中的规律；性繁殖动物利用其潜在伴侣行为的规律；社会物种的成员利用其同种个体行为中的规律；等等。即便是人类，这个似乎充满无限好奇心、储存大量可能永远用不到的信息的物种，也忽视了周围环境中的许多规律。举例来说，与积尘相比，你很可能在蚊子的行为中发现更多规律，即便你身边有更多积尘而非蚊子。如果你对蚊子叮咬有免疫力而对积尘过敏，可能就是另外一种情况了。

“相关推论必须利用实证规律”这一事实当然与经典推论方式不矛盾。经典推论方式依赖适用表征的形式程序或通用推论规则。在此基础上利用实证规律的方式便是对其进行表征，并把这些规律表征作为推论中的大前提。“如果……就……”式的陈述（比如，“如果这是一条蛇，就很危险”）就提供了一个简单的格式，可以用来表现许多规律，并将这些一般性表征与对特定事实的表征（如“这是一条蛇”）结合起来。形式规则可以从这个一般和特殊（或大或小）前提的结合中，推导出相关结论（比如，通过所谓演绎推理规则会推导出“这很危险”这一结论）。又或者，不仅一些而是所有的规律都能够用概率论术语来进行表现，而

且概率论推论规则可以接着应用于这些表征。

借助一小套正式推论规则，来充分利用庞大的规律表征及特定事实表征数据库，有助于形成一个正式且强有力的推论系统。可以说，任何无论怎样都可推断的事物，都可以用这种方式进行推断。不过，不要以为这种力量及普遍性有助于形成一个最佳的（乃至优越的）、自然选择应该会青睐的推论系统。

用逻辑性或概率论方式来推论具有普遍适用性，其替代性选择可为：利用许多专用模块，每个模块都借助适用任务的程序来利用给定规律。[注]举例来说，这可能就是某个物种自然而然害怕蛇类（无论是天生的还是后天的）时所发生的事情。专用推论程序把对周围环境中蛇的知觉作为输入信息，并将害怕的反应（通过其认知与行为层面）作为输出信息。这样一个程序既不依赖描述“蛇很危险”这一规律的前提，也不依赖假言推论的形式规则。发现蛇时就会立刻做出正确的反应，没有发现就不会有任何反应。

利用规律的程序并非像变魔术一样突然出现在进化或认知发展中，它们是对其所利用规律的存在与相关性的生物适应或认知适应。在这个意义上，这些程序包含有关规律的信息（正如钥匙包含所开启的锁的信息，或是抗体包含所中和的抗原信息一样）。

如果表征与程序不知为何都包含有关规律的信息，那么在规律表征与直接利用规律的程序之间又有什么区别呢？下面将揭晓答案。规律表征自身什么也不做，却能够提供为各种推论程序所利用的前提。专用程序会做的是：鉴于合适的输入信息，会制造出推论性输出信息。专用程序不会做的是：使其利用的信息能为其他程序所用。因此，举例来说，如果你有两种表征，“如果这是一条蛇，就会很危险”和“如果这是一只蝎子，就会很危险”，那么形式规则可能会让你推断出“蛇和蝎子很危险”或是“至少有两种危险动物”。另外，你可能拥有两种危险探测程

序，一种针对蛇，另一种针对蝎子，又无法将这两者合并起来做出这样简单的推论。

一个认知系统可以两次都包含同样的信息：一次是在直接利用该信息的程序中，另一次是在作为其他类型程序前提的表征中——你可能既对蛇有反射性恐惧，又知道蛇很危险。

通过具体程序利用规律和通过表征利用规律这两种方式，哪种更好呢？这么说吧，这个问题其实毫无意义。哪种方式更好，取决于可能因生物、环境、情境及意图而异的代价与好处。若生物的意图是要躲避蛇的伤害，那么一个迅速的、类似反射的专用模块很可能就是最好的选择。若生物的意图是要获取有关蛇类的常识，那么或许就应该运用类似陈述的一般性表征及更正式的论据形式。

没有证据表明，其他动物对常识的任何形式感兴趣（但我们也应该欣然接纳各种可能性）。就人类而言，所有人必定对躲避蛇类伤害（及其他类型的、有实际意义的具体知识）十分感兴趣，大多数人也对蛇类的某些常识感兴趣，而没有及时关注其实际意义。他们不仅想要利用规律，还想要对规律进行表征。这是否意味着人类只利用经典方式更好？还是两种方式都用？还是像我们指出的那样，其本身只是像经典方式的事物在人类认知系统中模块化了？

在关于模块性的论战中，有许多相关论据声称会证明，人类推论是基本经典化或模块化的。虽然让我们更加动摇的是支撑模块化观点的论据（我们也贡献了自身论据），[注]但是我们强烈地感觉到，若让两种替代性表述的概述一争高下，辩论起来会很痛苦。

经典方式发挥作用的时间要长得多，因此，不论是从形式角度来看，还是从实证角度来看，发展得都更加完备。然而，经典方式图景中仍十分简略（虽说不上有问题）的是，其解释的方式或解释不了的方式——人类推理在物种历史中是如何进化的、在个体中是如何发展的、是

如何成功地只进行对给定情境有重大关系的推论而未进行其所能进行的所有通常无关紧要的推论的。（这是并未出现的所谓框架问题，或至少在模块化系统中不是同一程度。）从最好的角度来看，仍比较粗略的也是经典图景解释的方式，这回是关于从同一前提出发进行推理的人为何经常会得出相异甚至相矛盾的结论。

我们计划引发这场辩论的方式不是把经典方式改头换面，而是充实模块化图景，特别是解释人类理性是如何适应这一模块化图景的。

---

1. Smith 2001.
2. 举例来说，Oaksford & Chater 2007; Tenenbaum et al. 2011。
3. Needham & Baillargeon 1993。雷尼·柏拉吉恩（Renée Baillargeon）为我们的初始图表提供了一个更高版本，对此善举我们非常感激。
4. Hespos & Baillargeon 2006；也可参见：Luo, Kaufman & Baillargeon 2009。
5. Fodor 1981, p.121.
6. Winograd 1975.
7. Fred Dretske (1981, 1997)、Pierre Jacob (1997), 以及Ruth Millikan (1987, 2004)等人的研究主张将表征作为广泛的自然主义概念，而这也正是我们使用的概念。若想获悉表征更严格意义上的视角，在解决哲学问题的同时也需关注经验主义证据，可参见：Burge 2010。
8. 我们并非在这里对表征在形而上学和因果关系（如果存在因果关系的话）中的任何特定视角做出承诺。可参见Egan 2012找到相关回顾和有所帮助的讨论。
9. 巴阿斯（Baars）与戴哈尼（Dehaene）的意识理论中有另一种假设的可能性，某些信息的传递并不是通过直接的模块与模块之间的连接来完成的，而是将一个模块的结果输出到其他模块可以获取信息的“整体工作空间”中(Baars 1993; Dehaene 2014)。
10. Tenenbaum, Griffiths & Kemp 2006.
11. Gallistel & Gibbon 2000；Rescorla 1988.
12. 由丹尼尔·丹尼特（Daniel Dennett）在很久以前提出 (1971)。
13. 例如，Sperber 2005。

第六章
元表征

心智真的只是许多模块的联合体吗？批评家认为，动物心智可能如此，但人类的一定不是！动物推论可能专门由只利用规律而无表征的模块执行。而人类却不仅能够利用，也能够表现许多实证规律。世界上的规律不仅仅是人类能够利用的东西，其他动物也利用规律，这些规律还是人类思考并谈论的东西。而且，人类能够有意识地利用实证规律的表征来发现更多的一般性规律。我们并非在为此而争吵。怎么能够如此呢？毕竟，正是通过练习这项能力，我们科学家才能够谋得生存。

更普遍来说，不正是推理的存在表明人类能够超越基于模块的直觉推论吗？最重要的是，理性难道不是从这些专用推论模块中分离出来另立门户了吗？不要这么肯定。我们将证明，推理是直觉推论的一种形式。

直觉与推理之间的经典对照，并不比动物与人类之间陈旧迂腐的对照更合理（而且人类拥有理性，野兽没有）。

不将人类与其他非人类动物相比，而是直接将人类与动物整体相比，这便使人类失去了正确理解以下两个问题的基本条件：“何为人”“人类如何从其他动物中脱颖而出”。同样，将理性与一般性直觉推论相对比，而不是与其他形式的直觉推论相对比，便是丢掉了理解人类如何及为何推理的方法。

大众本体论

如果理性基于直觉推论，那么你可能会问，直觉是关于什么的呢？我们将在第七、八、九章中阐述，运用理性时涉及的直觉是对理由的直觉。但首先我们需要做好准备。

对理由的直觉属于一个更宽泛的类别：对表征的直觉。轻松地对表征进行表征，并对这些表征进行各种直觉推论的能力很可能是人类心智最原始、最典型的特征。本章我们将着眼于这些对表征的直觉。

人类有非常丰富的“大众本体论”（folk ontology）。就是说，他们能够识别并区分世界上许多不同的基本类型的事物，而且是出于直觉，这是常识。大众本体论与科学本体论形成对照，大部分科学本体论既不是直觉性的，也根本不符合常识。随着年龄的增长，人的大众本体论也在切身经历与文化输入的共同影响下得到充实及修正，甚至还可能会被科学或哲学理论影响。然而，人类做的最基本的本体论区分在所有文化中都很常见（而且确定无疑的是，其中有些区分是由其他动物做出的）。

无论在何地，人类都能识别岩石这类无生命物体、鸟这类有生命物体、水与肉这类物质、颜色与重量这类物理属性、风暴与出生这类事件、进食与奔跑这类动作、勇气与耐心这类道德品质、数量或相似点这类抽象属性。一般来说，人类拥有对其在大众本体论中区分出来的许多不同事物的独特直觉。这就表明——且有充分的证据表明——人类在某种程度上拥有与不同的本体论类别相对应的不同推论机制。[注]

我们已证明，若推论中有可利用的规律，模块就可能会进化或发展——尤其是利用规律具有适应性的时候。其中许多规律都对应于本体论类别。举例来说，有生命物体与无生命物体以颇为不同的方式进行运动，并且其运动通常会给人类及其他动物构成非常不同的风险与机会。

存在一种相应的进化能力，能够识别这两类运动并予以区别对待。

然而，某些相关的规律，与本体论类别的基本特性关联更少，而与人类（或其他动物）的实际兴趣关联更多。举例来说，各种杂食动物，包括人类在内，可能会有对植物可食性进行推论的专用模块，虽然可食植物不是合适的本体论类别。事实上，模块对于任务、问题或机会的专门适用性，与其对领域的专门适用性一样频繁，如果不是更频繁的话。尽管如此，本体论仍是推论模块通常会利用的领域。

人类不仅对脑海及话语中许多种类的事物进行表征，也认识到自己此举在人类基本的本体论中（其中人类似乎很特殊），不仅有事物，也有对事物的表征。事实上，对人类能够表征的大部分事物，人类也能够对这些事物的表征进行表征。人类能够表征岩石、岩石的概念、颜色、颜色词、数字、数词、事情状态（比如，在下雨），也能对这些事情状态的表征（心里想着或说出“在下雨”）进行表征。

事物的表征本身就是世界上非常特别的一类事物。表征被视为特殊的（有子类别的）本体论类别，人类对此有专门的推论机制。对表征的表征，也被称为“高阶表征”或“元表征”，在人类认知及社会生活中起到了独一无二的作用。[注]然而，除了哲学家及心理学家以外，人们很少考虑或是谈及这类表征，而是在谈论表征的具体类型。

人们谈论信念、观点、希望、疑虑、恐惧、渴望或意图——所有这些都发生在人们的心智与大脑中——它们是心理表征，或者去谈论这些心理表征的公共表达，口头或书面语言、手势、图片——它们是公共表征。

心理及公共表征是位于不同时空中的具体事物：信念是某人大脑中某一时刻所持有的；口头陈述是发生在谈话者共有环境中的声音事件；书面陈述或图片是环境中的物体而非事件。然而，让这些心理及公共表征成为表征的不是其位置、持续时间或其他具体特征，它们更多的是作

为一种抽象特性，在常识心理学中被认为是“意思”或“内容”。当我们说起自己与某人有共同信念时，要表达的意思是我们所持信念有高度相似的内容。当我们说某个人表达了自身想法，意思是这个人说的意思与其想的内容相一致。

通常情况下，人们在思考或谈论表征时只考虑内容，并且从表征中提取出更为具体的特性。人们可能会说一个想法是真实的、矛盾的、混乱的、意义深远的或有诗情画意的，特别的是，不论是思想还是陈述，他们都不会把该想法归于任何人。如果他们这么做的话，就是在讨论抽象意义上的表征（简而言之，就是“抽象表征”）。像《小红帽》（*Little Red Riding Hood*）、《黄金法则》（*Golden Rule*）或乘法表这类文化表征在大部分时候都是抽象意义上的，它们必须在心理及公共表征中实现具体化，才能在人类事务中起到一定作用。

既然表征在我们的常识本体论中已得到认可，那么问题来了：若有的话，我们有何认知机制，可以对这些表征进行推论？对表征我们有何种直觉？我们也看到，存在好几种表征，而且每种表征都有不同的特性。没有先验理由来认定，人类有对一般性表征进行推论的模块。这种模块可能会利用何种规律尚不明确。另外，各种类型的表征呈现出可以利用的规律，用来进行不同类型的、具有高度相关性的推论。

婴幼儿、毛毛虫及一块隐藏的奶酪

元表征，即对表征的表征，由于大卫·普里马克（David Premack）与盖伊·伍德拉夫（Guy Woodruff）在1978年一篇著名文章中的提问，成为心理学研究的主要话题，该提问是：“黑猩猩有心智理论吗？”注。短语“心智理论”有助于形成一个吸引眼球的标题，也成为理论性误解的来源。普里马克与伍德拉夫的问题事实上并不是关于黑猩猩是否拥有对心智的理论信念（我们认为，或许是把对心理学规律的表征用作推论的

前提）。这个问题其实是，黑猩猩是否能够把具体的信念或意图归于彼此（或人类）。

有些作者，如艾莉森·高普尼克（Alison Gopnik）、亨利·威尔曼（Henry Wellman）和约瑟夫·佩纳（Josef Perner），认为要将心理状态归于他人需要有一些对心理状态的理论性理解，[注]还有些作者并不这么认为（我们与其想法一致），如雷尼·柏拉杰恩及艾伦·莱斯利（Alan Leslie）。[注]为了避免“心智理论”这一短语造成困惑，我们将使用“读心”（mindreading）这一隐喻来描述相关的认知能力。

人类能够“读心”，这一点再明显不过了。我们始终把心理表征归于彼此。我们常常能意识到周围人在想些什么，甚至能意识到他们以为我们在想些什么。我们很自然就能产生对他人想法的这类看法。

另外，没有证据表明，大多数动物，如沙漠行军蚁、蛇或牛，会把心理状态归于他人。牛可能在其本体论中并无心理状态。它们将其他牛视为具有生物学意义的行为方式的活体，如进食、反刍、睡觉、行进等，而不是视为执行基于自身渴望及信念的决定的执行者。对于其他少数尤为聪明的社会物种，如黑猩猩、狗、乌鸦或海豚，普里马克和伍德拉夫所提问题仍具有争议性：是的，这些动物或许能够进行一些初级的“读心”，但远不及人类在此方面的能力。

普里马克与伍德拉夫的文章对动物心理学及儿童心理学研究都产生了巨大的影响。儿童从几岁开始解读心智呢？该领域中所有研究都起源于20世纪80年代早期，研究表明，在4岁左右，儿童便能够轻松地将错误的信念归于他人（不管是好是坏，都已成为检验真正的“读心”的“试金石”）。[注]接着，欧尼士（Onishi）和柏拉杰恩于2005年进行的开创性研究[注]表明，不仅4岁的孩童，甚至婴幼儿也会对周围人的脑海中在想些什么有所关注，甚至还期待执行者的行为与其信念相一致，不论其信念正确与否。随后更多研究证实了二人的发现。

举例来说，卢卡·苏里安（Luca Surian）、斯坦法纳·卡尔迪（Stefana Caldi）和丹·斯珀伯（Dan Sperber）向13个月大的婴幼儿展示了一段“好饿的毛毛虫”视频。㊟这条毛毛虫看到一只手在其左侧的屏风后面放了一块奶酪，又在其右侧的屏风后面放了一个苹果，它爬向左侧屏风并开始啃咬奶酪（见图6.1）。这些孩子看了好几遍这个视频，所以很清楚这条毛毛虫更喜欢吃奶酪。接着就是关键性的测试阶段了。这次，孩子看到那只手将奶酪放到了右侧屏风后，将苹果放到了左侧屏风后。孩子看到毛毛虫是在两样食物都被藏起来之后才到达现场的，所以认为毛毛虫不会知道食物的位置已经调换。

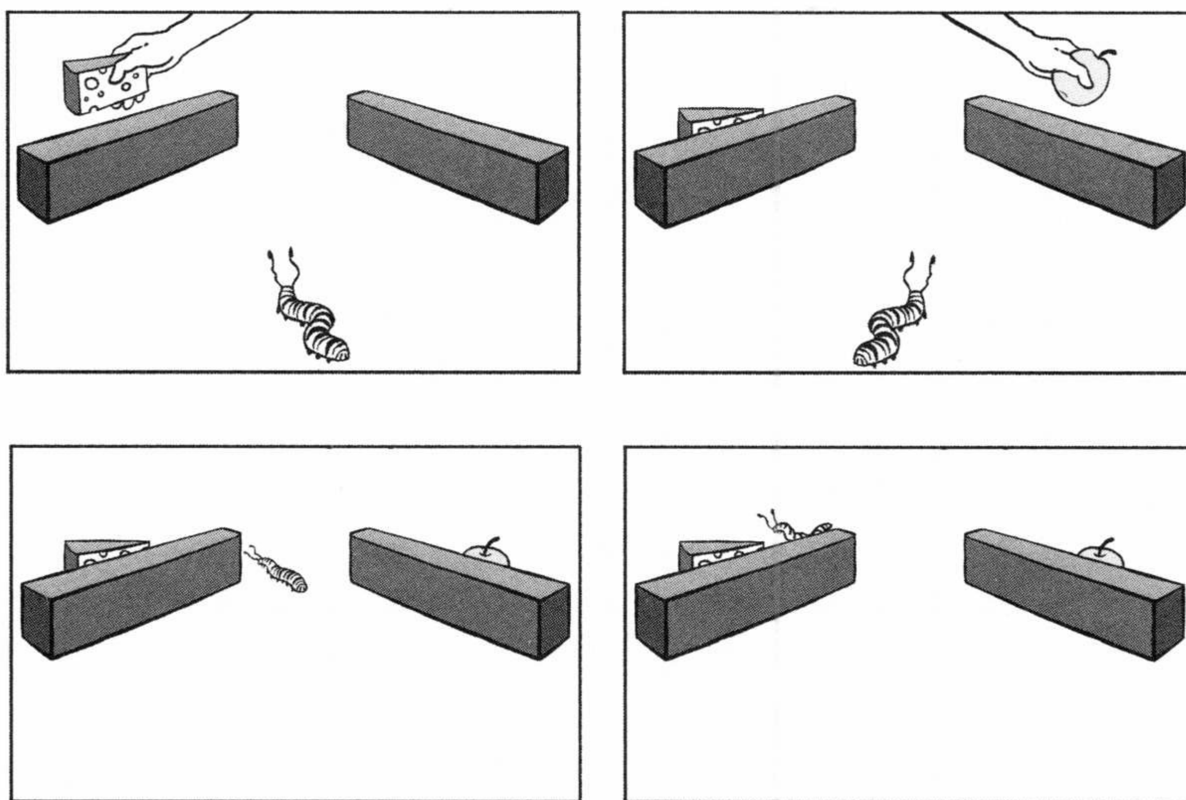


图6.1 在熟悉阶段，婴幼儿看见毛毛虫爬去啃咬奶酪

在这种情况下，孩子认为毛毛虫会怎么做呢？是去它屡次看到藏着奶酪的左侧屏风后，还是去这次藏着奶酪的右侧屏风后呢？作为成人，我们认为毛毛虫很明显会去它相信的奶酪会在的地方，即左侧屏风后，即便这种信念在此情况下是错误的。但孩子能够做出这种推论吗，特别

是把错误信念归于他人？令人称奇的是，婴幼儿的表现与成人一样。他们也料想毛毛虫会去左侧屏风后，因为毛毛虫有理由相信（错误地）会在那里找到奶酪。当毛毛虫径直爬向右侧屏风的奶酪时，孩子看的时间更长，因为毛毛虫无从知晓奶酪的位置。也就是说，婴幼儿期待执行者根据自身信念行事，无论这一信念正确与否。

我们是否可以得出这一结论：该研究中的婴幼儿对“代理人理性地形成信念及意图”这一普遍心理事实具有心理表征，而代理人将这一心理事实作为推论的前提？再一次，这样想更有意义：婴幼儿拥有一个专用程序，该程序用执行者形成信念及意图的方式来利用某一规律。婴幼儿并不需要对这一规律进行表征。

“读心”要如何利用“人类和毛毛虫这类执行者往往是理性的”这一事实，而不把对这一事实的心理表征用作前提呢？我们在第三章至第五章中已经表明，使执行者变得理性的不是一般性机制，或理性地思考与行动的倾向，而是有专门用于不同推论的各种推论机制。尽管这些机制五花八门，却都受到了有助于提高效率的选择性压力的影响——在这种情况下，这些机制的功能便在于用高效的方式进行这种推论。

可以说，“理性”从最基本的意义来看，与推论效率是同义的。能达到怎样的理性程度，在很大程度上取决于众多推论模块各自发挥功能的方式及这些模块的连接方式。为了利用执行者的理性，“读心”就必须求助于具体的推论机制，并充分利用这些机制进行专门、有效推论的倾向。

为更好地理解真实的人这种多方面的理性是如何能被“读心”利用的，我们最好走出实验室，避免只关注“错误信念任务”及少数相关的实验范例，不论它们设计得多么好。

下面是一个关于日常生活中“读心”的简单例子。当你走进医生的候诊室时，已经有个患者在等了。你俩都瞥了对方一眼并说了句“你好”。

你坐了下来。她偶尔在自己的智能手机上打字，并盯着手机屏幕。你取来一本杂志。她看了看自己的手表叹了口气。你们又互相瞥了一眼。毫无疑问的是，你们都有各自的一连串想法，这些想法对方并不知晓，但尽管如此，你们两人都自然地进行了些微交互性“读心”。

当你到达候诊室时，她知道你明白她会比你先看医生；看到已经有人在等待了，你感到很失望，但你试着不去表露出来，不让她看出你的失望情绪，但她可能还是看出来了；你知道，她时不时地点击自己的智能手机并一直盯着屏幕是在与他人保持互动，而这个与她聊天的人可能跟她有一定距离（即便你不知道他们在聊些什么）；你猜她看手表是因为那时候早已过了她与医生约定好的时间，而她在叹气则是因为她不喜欢等待；从相互一瞥中你明白，她知道你领会了她的叹气声；等等。所有这类发生在最低限度互动中的“读心”，都是自然而然进行的，毫不费力。

同类之间即便不存在互动也能进行本能性“读心”，即在一个人只是看着另一个没有意识到自己正在被看的行为人时。13个月大的婴幼儿推断，毛毛虫已经屡次看到奶酪被放在左侧屏风后，所以这次还是会以为奶酪在同一个位置。不像毛毛虫，婴幼儿已经看到奶酪这一次被放在了右侧屏风后面，尽管如此，他们还是认为毛毛虫会去左侧屏风后面寻找奶酪，就像之前一样。

这类“读心”是多么明显啊！然而，不明显的是，是什么使这类“读心”变得如此明显。下面是对所发生之事的概述。我们人类倾向于对自身社会环境进行持续监控（正如在候诊室中的例子一样）。我们打开、维护并更新关于自己认识的所有人的“心理文档”[注]（包括一些我们只是听说过的人，如国王或著名演员，更不必说那些我们很了解的科幻人物），在这些关于人（及其他种类的执行者，如毛毛虫、诸神）的“文档”中，可能会存在各种各样的信息：关于其姓名、家庭、历史、外貌、性情、举动的信息；也有关于他们储存他人信息的“心理文档”中的

信息，其中的他人也包括我们自己。我们关于他人的“心理文档”包含关于他们“心理文档”内容的信息，而且那一信息是由“读心”提供的（当然也是元表征性的）。

其中一些关于人的“心理文档”是非常单薄且短暂的，就如你在医生候诊室中对另一个患者打开的“心理文档”一样。而其他“心理文档”则厚实且持久，就如关于你亲密的家庭成员的“文档”。这些“文档”中部分“读心”信息来源于你在观察他人举动时的自发性解读。部分信息则是由人们自己提供的，这些人在与你沟通时，帮助你解读其自身心智。更多信息则是通过人与人之间的互相谈论而来：八卦。关键是，我们是基于许多不同的证据、通过许多不同的方式来解读心智的。

一定存在一个“读心”模块——事实上是一个能够解读许多心智的模块——要做这样一个工作，那就是在我们关于他人的“心理文档”中，整理他人在其自身“心理文档”中的东西。然而并没有可以独自完成这种工作的模块。为了能够对在他人心智中进行的推论进行“读心”式推论，“读心”模块必须与其他许多不同的推论模块建立连接，并且使用这些不同的推论模块来对表征在个人“文档”中的信息进行更新。

举个非常简单的例子。蒂姆（Tim）问你要糖果，于是你给了他一些。他看着糖果说：“就三颗糖？嗯？”并把这些糖放进了一个空纸袋。在你保留的关于蒂姆的“文档”中，你认为他的信念是“纸袋子里有三颗糖”，现在就有了你对该信念的元表征。他说：“再多点吧，好不好？”你又给了他一小把糖，他又看了看糖，说：“五颗糖，谢谢，加上第一回你给我的那三颗，现在够了。”然后又把它们放到了纸袋子里。在你关于蒂姆的“文档”中，蒂姆的信念是“纸袋子里又多了五颗糖”，那么你现在就有了对该信念的元表征。碰巧，你数出来的不是五颗糖而是六颗糖，所以你认为蒂姆的信念是错误的。

蒂姆认为他先是往纸袋子里放了三颗糖，接着又放了五颗糖，基于他的这一信念，你可能会把“袋子里有八颗糖”这个进一步（错误）的信

念归于他。但你是怎么进行这种“读心”式推论的呢？进行数学运算——举例来说，3+5=8——并不是你的“读心”模块为自身设置的运行方式。为此，你的“读心”模块必须将蒂姆的两个相关信念的内容（他第一次往袋子里放了三颗糖，接着又放了五颗糖）与你的数学模块共享，并且将其运算结果拷贝出来，放入你关于蒂姆的“文档”中，然后就能对蒂姆认为袋子里有八颗糖的信念进行元表征了。鉴于你本人的信念是“六颗糖加到了最初的三颗糖中”（而不是蒂姆误认为的五颗），因此你的数学模块便得出结论，现在袋子里有九颗糖。然而，这却输入了你为那个袋子打开的“心理文档”中（我们不仅对人有“心理文档”，而且对各种各样的事物都会有“心理文档”），而不是你为蒂姆打开的那个“文档”里。

事实上，越来越多的证据表明，我们高度社会化的心智一直追踪与预测的，不仅是自己周围物理环境中发生的事情，还有在周围人心智中发生的事情。[注]为了实现这种双重追踪，我们的许多推论模块都在惯常地进行推论，不仅是为了更新我们自身对周围世界的信念，也是为了更新我们对于周围人信念的元表征。

这就意味着我们是通过用他人的观点来思考从而知道他们在想什么的吗？事实上，只有完整的人才会有观点，才可以尝试用他人的观点来看事情。另外，个体大脑中的心智模块是“次人格性”（subpersonal）[注]机制，并没有观点。

模块会偶尔使用“离线”（offline）来模拟他人的心理过程吗？我们更倾向于认为，追踪他人的心理过程是其日常“在线”（online）工作之一。举例来说，我们的数学模块是以中立的立场来进行数量运算的，而这些运算的输出则可能会更新我们对于事物情况的表征，同时也会更新对他人心智中表征的事物的元表征。[注]

在毛毛虫实验中，婴幼儿屡次看到奶酪放到左侧屏风后，也看到毛毛虫看到了同样的事情。因此，婴幼儿为毛毛虫及自身进行推论，左侧

屏风后就是奶酪通常会在的地方。毛毛虫不在场的时候，婴幼儿看到奶酪放到了右侧屏风后，因此他们对于奶酪所在位置的表征已经得到更新，而其对于毛毛虫表征的元表征还未得到更新。接着，当毛毛虫到达现场时，婴幼儿的“读心”模块使用的是元表征在毛毛虫“文档”中的信念（这个信念现在错了），来预测毛毛虫会爬向哪里寻找奶酪。

那么，婴幼儿又是如何建构有关毛毛虫未来行动的合理期待的呢？我们认为，在毛毛虫到达现场的时候，婴幼儿的“读心”模块：（1）将信息从毛毛虫“文档”转换到目标导向运动模块，该模块的任务是在空间中计算到达目的地的合理路径；（2）使用这一计算结果来更新毛毛虫“文档”并预测其行动。

在候诊室的那个例子中，你从众多模块中选用了一种模块化能力，这种能力受到社会规范的影响，这种能力能为你在以下情境中提供指导：你的身体碰巧暂时靠近陌生人，因为你们各行其道而非目标一致（如在候诊室里、电梯里或飞机上）。在候诊室中，你从个人视角及另一个病患的视角，注意到了与你的互动相关的事件。你解释道，自己简短的致意、短暂的一瞥，只是你维持你俩可能都感到最舒适的社会距离的方式而已。为了进行这种“读心”，你不必积极地决定采用另一个人的视角。你从一开始就自动更新了你打开的关于她心理状态的“文档”，正如你以更为精细的方式来更新自己对于物体及周围事件的“文档”一样。

我们认为，这种对于周围其他人心理状态的自动化、持续性追踪，常常涉及信念、意图、决定及其他实质性心理状态属性。它自从婴幼儿时期就开始发展，直到成人阶段（存在个体及文化差异），已发展成相当完备精细的、迅速理解他人的能力。注

文化传统以多种方式增强、阻碍并影响我们彼此理解的方式，社会角色及职业分工也是如此。在不同的文化与历史时期，即便是在同一文化内部，人们对人类心智也有许多不同的明确看法，这些看法能够通过

箴言、精致的民间或学术理论来传达。然而，我们每天在与他人互动时（或是在听到有关他们的话语、看到有关他们的文字时）所进行的“读心”仍然十分自然，并颇为直觉化。“读心”不会将这些文化观念用作旨在识别他人心理状态的自发性推论的前提。这些观念是在需要时用来帮助解释及佐证得出的颇为直觉性的结论的。

虚拟领域

“读心”为我们提供了关于他人信念、渴望与意图的直觉，以及这些因素如何与人们的感知和行为相关的直觉，绝不会耗尽我们对表征进行推论的能力。举例来说，数字认知就为我们提供了一种截然不同的元表征性直觉。

最近许多基于数字认知的研究工作已经表明，[注]人类婴幼儿与其他动物都有对分散物体的数量进行心理表征的能力。对于小数目（如1、2、3）而言，这些表征非常精确，对于大数目则是粗略估计。习得数字名称让人类有了词汇工具，可以用精确的方式来表征比“3”大得多的数目。与语言及书写相联系的、从文化中萌生的数字系统（即符号），可以对数量进行公共表征，在某些文化中，已经促进了全新的知识分支——算术的发展。任何学过一点算术的人（即便看起来每个人都只是在用数字数数）[注]，都不仅有对具体数量的直觉，还有对数字间形式关系的直觉，这些关系是通过表征它们的数字符号而被理解的。

只举一个例子来说明，对你而言，900是300的3倍在直觉上颇为明显，你不需要运算就能知道。这个直觉利用了你对极小数目间关系的知识（在这个例子中，3×3=9），还利用了表征这些关系的特殊方式——十进制（用阿拉伯数字）的属性，而非数量属性。如果你使用的是九进制，那么1 210是363的3倍多看起来就不那么明显（即便这两个数字表征的是十进制中的900与300）。另外，如果你使用九进制，那么1 000

是300的3倍多对你而言就颇为明显。这一次，反倒是在十进制中，729是243的3倍则不是那么明显（即便这两个数字表征的就是九进制中的1 000与300）。我们对于整数有更好的直觉；然而，凑整却不是数的属性，而是用以对其表征的数字系统的属性。

这个关于数与数字的例子阐述了以下三个关键点。

1. 我们对事物（这里指数）的直觉与我们对其表征（这里指数字）的直觉并不相同。

2. 我们对表征的直觉利用了表征的特性，然而所表征之物的特性（如凑整）不一定与之相一致。

3. 尽管如此，我们对事物表征的直觉可能是对所表征之物本身的顿悟来源。（例如，900是300的3倍，这是关于数本身的事实；这个事实能从直觉上被领会，是因为十进制中用来表征这两个数的数字间存在直觉性关系。）

数是一种很特别的事物，而数字则是一种很特别的表征。尽管如此，我们刚才所列举的关于其关系的三个通用性关键点，还是能够很好地延伸到其他类别的元表征性能力中的。

拿解释来举个例子。解释是某种类型的表征。儿童在远远不能自行给出任何类似的真正解释之前，会提出各种各样的问题。这说明，他们开始索要对大量不同事物的解释。很快，他们就开始给出自己对事物的解释了。更宽泛地说，在不同文化之间，索要与给出解释是交谈中非常普遍的一个方面。

对不同解释的“好处”，我们有清楚的直觉。正如心理学家弗兰克·凯尔及其同事表明的那样，如果我们给出解释的能力遭到质疑，这些直觉或许就不太可信了。[注]举例来说，我们常常严重高估自己在解释我们

每天使用的家用电器实际是如何工作方面的能力。然而，我们却能更好地评估别人给出的解释。即便是孩子，通常也能够颇为熟练地发现别人在给出特定领域的解释时具备专业性，并且很擅长对其加以利用。凯尔描述了不同领域中解释寻求者与提供者的认知劳动分工，其与交流观点时通常发生的认知劳动分工颇为类似（我们将在第十五章中进行展示）。

以下几点与数及数字的例子完全相同的是：

1. 我们对好的解释与坏的解释的直觉，与我们对所解释之物本身的直觉是不同的。

2. 我们对解释的直觉需要用到的特性，如说服力、通用性或连贯性是解释本身的特性，而非所解释之物的特性。

3. 尽管如此，我们对解释的直觉（会让我们更偏爱好的解释），仍是对所解释之物主要的顿悟来源。

更宽泛地说，表征是世界上一种非常特殊的东西，只在心智内部或是邻近的地方才能够发现其存在。按照合理的评判标准，元表征性模块及其对表征的推论，都是高度专门化的、针对特定领域及任务的装置。与此同时，我们对表征的推论与对所表征之物的理解高度相关。

如果你凭直觉知道900是300的3倍，那么你的直觉便是受到数字特性的驱使，但是你获取的相关信息是关于这些数字所表征的数。如果你认可一个合理解释的说服力，如关于双重科技运动探测器如何运作的解释，那么你学到的就不仅是关于这个解释的知识，更重要的是，还有关于运动探测器的知识。

可以说，“读心”是我们元表征性模块最为重要的部分，这一点毫无疑问。“读心”让你通过对他人所思所想及信念主题的直觉来获悉事情。

当候诊室中的女士看着自己的手表叹了口气时，你认为不但是她觉得医生在让她等着，而且事实上医生确实是在让她等着。从这一点，或许你就可以继续推进自己的推论，如这个医生是如何对待自己的预约的。通过发现他人对这个世界的想法，我们可以学习到有关这个世界的许多知识。“读心”为我们打开了一扇详尽观察这个世界的窗户。

元表征性模块提供的信息不仅关于被元表征的表征，还间接性地与这些表征所表征之物相关。因此，我们认为，这些模块有高度专门化的领域，即具体种类表征所对应的具体方面，尽管如此，它们可能还有一个与其所表征之物对应的、不同的、更广泛的虚拟领域。

对人类心智的模块化观点的一般反对意见是，推理既然不是专门化的、局限于单一领域的，就不可能是模块化的。确实，人类能够而且也确实在对所有事物、任何事情进行推理。人类还将其关于几个不同领域的推理证据统一到一起。

举例来说，为了计算出地球周长，亚历山大港学者埃拉托色尼结合天文学、地理学、几何学知识，对阴影进行实地观测。或者拿一个更近一点的辩论作为例子，即以“都灵裹尸布”（Shroud of Turin）著称的那块古代亚麻布。它是像有些人认为的，是圣徒耶稣的裹尸布并因此属于圣迹，还是像另一些人否认的那样呢？关于该问题的论据涉及神学、化学的各个分支及物理学（特别是放射性碳追溯技术）。这两个例子——埃拉托色尼的发现与关于都灵裹尸布的辩论——完美地阐释了人类推理并不局限于某一特定领域吗？更不用说模块化了。

事实上，这些阐释在催生出有趣问题的同时，并不对模块化理论提出强烈的反对意见。这些阐释忽视了一个事实，那就是元表征性模块可能远远超越了其真实领域的虚拟领域。

我们将证明，推理是基于元表征性模块的，这一模块提供的直觉不仅关于整个世界，还关于理由。理由是一种表征。理性模块的真实领域

——理由——是较为狭隘的。尽管如此，理由本身可以是关于任何事或世界上任何事物的结合，举例来说，能够将沙漠中骆驼的步调与亚历山大港夏至日正午时分天空中的太阳位置统一到一起，能够将圣徒耶稣受难与放射性碳的放射性衰减比率联系起来。因此，推理在运作中保持相当专门化的同时，也能够在含义上保持相当的通用性。事实上，推理的普遍含义可以得到其专门领域的具体特性的最佳解释。对那些关于任何事物的理由进行推论会导致一种虚拟领域普适性。

---

1. 例如，可参见：Carey 2009；Hirschfeld & Gelman 1994; Sperber、Premack & Premack 1995。
2. 可参见：Sperber 2000。
3. Premack & Woodruff 1978.
4. Gopnik & Wellman 1992；Perner 1991.
5. Baillargeon, Scott & He 2010；Leslie 1987.
6. Baron-Cohen et al.1985；Wimmer & Perner 1983.
7. Onishi & Baillargeon 2005.
8. Surian, Caldi & Sperber 2007.
9. 对于“心理文档”概念的多样化运用，可参见：Kovács 2016; Perner, Huemer & Leahy 2015; Recanati 2012。我们的方法更接近于Kovács的研究。
10. Frith & Frith 2012; Kovács, Téglás & Endress 2010; Samson et al.2010.
11. 人际水平与次人际（subpersonal）水平之间的区别可参见丹尼特（1969）做的介绍，有几种不同的解释，其中包括丹尼特自己的解释版本；可参见：Hornsby 2000。
12. 哲学家阿尔文·戈德曼（Alvin Goldman）与神经科学家维托瑞奥·盖乐瑟（Vitorrio Gallese）就一种颇有影响力的学说进行过专门的辩论，是指当我们对他人心理过程进行模拟时所理解的他人头脑中进行的活动（如Gallese Goldman 1998)。我们认为，这种方式会随着“模拟理论”的标准版本而变化，并且与其他版本更加匹配；但在其他版本中，“模拟”是在技术意义上的“模拟”，比其自身原本的意义要更为宽泛，因此造成了许多误解。
13. 一些心理学家与哲学家（例如，Apperly & Butterfill 2009; Heyes & Frith 2014) 认为，追踪他人心理状态的机制发展尚未完全，连信仰、意图这样的心理状态也还无法察觉。据这些专家说，由“通过”标准错误信仰任务的4岁儿童［或是喜欢看简·奥斯汀（Jane Austen）小说的成年人］表现出来的更为精细的理解，是基于完全不同的“读心”“系统

二”机制。这种二元假说的主要优势在于将与婴幼儿“读心”有关的新近证据与“读心”早期标准视角的最小化修正相适应。另外一些人——我们也同样同意他们这一观点——则更偏爱对这些标准化视角进行修正，对支持以双重系统方式来理解“读心”过程的所谓证据提出疑问（可参见：Carruthers 2016; 也可参见：Brent Strickland & Pierre Jacob, “Why Reading Minds Is Not like Reading Words”,January 22, 2015, 网址：http://cognitionandculture.net/blog/pierre-jacobs-blog/why-reading-minds-is-not-like-reading-words)。

14. Dehaene 1999.

15. Gilmore, McCarthy & Spelke 2007.

16. Rozenblit & Keil 2002.

第三部分
反思理性

理性是什么？如何运作？是为了什么？如何进化而来？在第七到十章，我们运用新颖的互动理论来回答这些问题。我们认为，理性是关于理由的直觉推论机制，逻辑在这些理由中充其量只扮演一个不重要的角色。人类用理由来为自己辩护，并说服他人。在他们的合作与交流中，这两个活动发挥着关键作用。蝙蝠为适应生态位而进化出了回声定位能力，与此相似，人类利用错综复杂的社会关系、强有力的语言及丰富的文化为自己建造并维系十分特殊的生态位，而为了适应这种生态位，人类的理性也得到进化。

第七章
如何使用理由

人类不仅在推理中会用到理由，在解释及为自身辩护时也会用到。注尽管如此，解释和辩护是一方面，推理是另一方面，对两者的研究相对独立，仿佛是两个完全不同的心理学机制。我们认为，二者在功能上的区别基于在机制上的区别，而且，在本书的第三部分，我们想用一种综合的方法来清晰地呈现理由心理学。

你为何要这样想？你为何要那样做？我们借助理由来回答这类问题，仿佛理由指导我们的想法及行动并因此给出解释是不言而喻的。这些理由可以被公开评判：或好，或坏。好的理由为人们所解释的想法或行动辩护。理由在解释和辩护方面起的作用很明显，却是基于一个便利性假设：大多数的理由都是事后的合理化。尽管如此，人们依然在互动中充分利用理由的便利性假设，从最细枝末节的互动到最戏剧化的互动，都是如此。

常识图景

我们从戏剧性的互动开始探讨：2013年11月2日晚，中年男子西奥多·韦弗（Theodore Wafer）被前门巨大的爆炸声惊醒，他住在密歇根州底特律市近郊的迪尔伯恩。那晚早些时候，年轻的非裔美国妇女雷尼莎·麦克布莱德（Renisha McBride）的小轿车被撞毁，而后她从车内逃出来，在街上意识混乱地游荡了数个小时。最后她在韦弗的门口停了下

来，很可能是为了寻求帮助。韦弗以为自己的房子被攻击了，于是拿上猎枪，打开门，然后开枪杀害了麦克布莱德女士。在审判中，他解释此举只是因为生命受到威胁，他以为有好几个攻击者，开枪只是要自卫。控方认为他这是非理性的恐惧。不管怎样，即便他的恐惧是理性的，他的反应也应该是不开门，将自己锁在屋内，然后打电话叫警察。因此，韦弗被判二级谋杀罪。

为什么韦弗会有那种反应？无疑，他在审判中给的理由是他能想到的最好的回答了。我们永远也无法得出确切的定论。很有可能他也不知道自己那时到底是怎么想的。只是说，我们接受他基于这些理由做出的开枪辩解：他感觉正受到攻击，所以想要保护自己。但即便如此，我们大多数人还是会同意控方以及陪审团的意见，即他并没有足够合理的理由来认定自己处在巨大的危险当中，也没有足够合理的理由开枪。

正如韦弗的案例展现的那样，一些用来解释的理由可能看起来足够合理，但若用于辩护是不够的。我们能接受一个解释，但同时会批判性地看待其所援引的理由。一方面，为了达到解释的目的，只要我们试图去了解的这个人自认为这些理由足够充分就行了。另一方面，若要断定这个人的想法和行为合乎情理，他的理由一定要在我们看来也是充分的才行。

不用说得这么戏剧化，在日常的社会互动中，我们经常会援引次要甚至是微小的理由。譬如，罗博（Rob）询问智恩（Ji-Eun）：“你想要一杯奶昔吗？”智恩答道：“谢谢，但是你知道，我们韩国人大多数都受不了乳糖。”她为何不直接说不要呢？为何要费事地给出拒绝的理由？在婉拒罗博的好意时，智恩给出一个理由，是想表明她很感谢罗博的好意，要不是这一理由，她就会欣然接受。罗博不可能去质疑智恩援引的理由。对他而言，重要的应该是智恩给出了理由而不是只说了“不要”，此举表现了她对罗博的贴心考虑。在这种普通的互动中，我们给出理由就表现了一种入微的体贴，这是他人对我们的期望，也是我们对他人的

期望。

为了能像在思考和互动中那样使用理由，人们应该如何理解理由？要回答这个问题，只着眼于人们如何理解“理由”（reason）这个词本身（即做某事的理由这层意义而言），是没有多大用的。这一理解因人而异、因环境而异，显然也因文化而异。如此说来，英语单词“reason”在所有语言中都没有直接对应的翻译。总之，人们思考并谈论一般性理由时的这种模糊性与多变性本身就值得研究。我们这里研究的不是理由的民俗概念或者民俗理论，而是人们利用某些特定理由的方式，以及他们利用这些特定理由时是否把它们归为“理由”。

要更好地了解理由产生及运用的方式，区分客观理由与心理学理由将大有裨益。这种区分（用微妙的差异和不同的术语选择）在关于理由卷帙浩繁的哲学文献中很常见。我们为了达到当前的目的，会在必要时从这类文献中汲取灵感且不会卷入争论中。注

客观理由是客观地支撑某一结论的事实。这一结论可能是描述性的（是什么情况）或者是实践性的（什么是令人满意的）。例如，“今天是星期五”这个事实就是推出“明天是星期六”的客观理由；“梅子成熟”这个事实就是“要立刻把它们从树上摘下来”的客观理由。

我们用术语“事实”来指真命题和没有因果力的抽象客体。严格来说，导致梅子从树上落下的不是“有些梅子熟了”这个事实，而是“梅子熟了”这个现象本身，或者更准确地说，是定义“成熟”这个概念的化学及物理变化才让梅子从树上落下来。作为事实的客观理由本身并没有因果力。从我们的实证性观点来看，客观理由是否在人类利益交织的世界里独立存在、是否符合哪种定义明确的客观性标准都没有关系。就算我们真要谈论客观理由的话，也是因为它们在人类的思想和语言中被表征，而且，与事实不同，事实的表征的确是有因果力的。

心理学理由是对客观理由的心理表征。像所有的心理表征一样，心

理学理由可能是错误的表征；也就是说，作为支撑某个结论的客观理由，心理学理由可能会对错误的命题进行表征，而非对事实进行表征，或者说，它可能会对真正的事实进行表征，但支撑着自己其实并不支撑的结论。尽管如此，从认知及进化的视角来看，我们猜测，一般来说，人们往往会将真正支撑某个给定结论的事实，表征为支撑该给定结论的客观理由：认知虽不完美，但也并不随意。心理学理由是存在于大脑中的心理表征，并因此在人们的生活中发挥着因果性作用。（当我们不加限制地使用“理由”这个术语时，就是在谈论心理学理由。）

人们普遍认为，心理学理由发挥的主要作用是激励和指导人们的行动和信念（指导不过是激励的一种更精细的形式而已，而激励则是指导的一种粗糙化形式）。我们反对这种看法。理由所发挥的主要作用不是激励或指导我们得出结论，而是解释和辩护我们已经得出的结论。

理由的解释作用和辩护作用不仅相互关联，而且交织在一起。要解释人们的信念或决定，心理学理由最起码是可以相信的正当理由，即充分、客观的理由。韦弗或许没有足够的理智，但也不是荒谬至极，因为他能拿出理由来为自己枪杀门口到访者的举动辩护。但是，要是他给出的开枪理由是“猫王埃尔维斯·普里斯利（Elvis Presley）已死，因此生命变得毫无意义”，我们不会把这个理由当作真正基于理性的解释，甚至不会当作瑕疵性理由，而是作为暂时性精神错乱者的供认（或声称）。

同样，当我们援引理由来为他人的想法或行为辩护时，通常指的是他们将这些理由牢记于心，而且受其指导。人们如何用自己一无所知的理由来为自己辩护呢？这似乎是对辩护符合常理的约束，但这是真的吗？事实上，哲学文献中描述的“道德运气”或“认知运气”的案例便是有趣的例外。注

譬如，想象一下，如果韦弗射击并杀害的这个人不是一个寻求帮助的年轻女性，而是在“头号通缉犯”名单上出现的危险罪犯，那么又会是

怎样一番情形呢？那样，韦弗的所作所为可能会得到表扬，即使他是在不知道此举，碰巧有充分、客观的理由时而这样做的，这就是道德运气的典型案例。虽然人们必须想出理由来解释自己的行为举止，但至少在某些案例中，他们不需要想出理由来使自己的行为举止被谴责或表扬。

就如存在道德运气的例子——在脑海中没有充足理由时而采取了行动，结果却是，在行为人不知道的情况下，就已经存在非常客观的正当理由——也存在认知运气的例子：人在理由不充分时持有某个信念，结果却仍然合乎情理。

正如我们在第一章看到的，亚历山大港学者埃拉托色尼是测量地球周长的第一人。他的计算结果是252 000斯塔德，只与现代测量的24 859英里相差1%。对这一精准性，更加让人不可思议的是，埃拉托色尼在其设想中犯了两个错误。他的估算中极其重要的一点是，认为赛伊尼城正好在北回归线上、亚历山大港的正南方，他也相信情况确实如此。但实际上，赛伊尼城在北回归线偏北1°、亚历山大港偏东3°的地方。所以他是如何做到如此接近的呢？纯粹是运气，这两个误差的影响实际上相互抵消了。这一点认知运气并没有被用来贬低埃拉托色尼的成就。仿佛他的一些“理由”变得无关紧要了，因为他的全套方案太绝妙，成果太令人钦佩。

道德运气和认知运气是令人费解的现象：二者表明，我们可能会发现，有的理由或许并没有激励或指导人，却能证明人的正当性，因而人无法解释自身的想法或行为。理由的常识图景是不是不如它初看起来那么清晰且连贯呢？

遭质疑的常识图景

与常识图景相悖的是，许多实验证据表明，人们在形成信念或做出

决定时，常常不怎么关注，甚至根本不关注理由。人们使用理由主要不是用来指导自己，而是维护自己在他人眼中的正当性，并评估他人的正当性（通常是批判性的）理由的。当我们真的用理由来做指导时，大多数时候是用来指导他人而不是自己。虽然我们想让他人受到我们所给理由的指导，但最后总会让自己受到自身直觉的指导（虽然我们不能详细说明这些直觉，但确信它们是基于充分的理由）。

不管受到理由的指导是否会更好，事实是，为了相信或决定某事，我们不需要关注理由。我们的许多信念及决定都源自纯粹的直觉推论，其运作方式令人费解。你看着朋友莫莉，莫名的直觉让你觉得她很沮丧。对此直觉你有何理由支撑？又比如，你查看今夜剧场放映的是哪些电影，看到了《星球大战12》（*Star Wars 12*）和《超人8》(Superman 8)。你最后决定去看《超人8》。你做此选择的理由何在？当然，如果有人问你，你会编出一些，但事实上，当凭直觉知道莫莉很沮丧或选择《超人8》时，你是无意识的，更不用说仔细考虑了。你的想法和选择都是出于直觉。

可能仍然会有人说，理由已在无意识中指导我们了。他们还认为，我们通常会自省，从而会意识到我们无意识的理由。但真的如此吗？当我们为自身辩解时，我们真的能意识到那些在无意识中指导着我们的理由吗？我们先来看一些富有挑战性的证据，然后在下一节再提出更加彻底的疑问吧。

无疑，西格蒙德·弗洛伊德（Sigmund Freud，奥地利精神分析学家）的著作已渐渐摧毁了人们对心智解读能力的常识自信，其著作关注的是所谓的“无意识”。虽然托勒密和伊本·阿尔-海什木很久以前便已承认了无意识心理过程的存在，但是直到弗洛伊德之前，这些过程都被视为是相对边缘化的。心理活动多半时候都是典型有意识的状态，至少是可以内省的。但是，弗洛伊德给出的案例令人信服，让人相信他们相当普遍地弄错了自身真正的动机。100年过后，从认知心理学来看，曾经

极具挑战性的“无意识”观念似乎过时了。不是一些，而是所有的心理过程，包括情感的和认知的，现在都让人觉得很大程度上甚至完全无意识。如果有问题存在，那么问题已经演变成为何存在及质疑怎样存在意识这种东西。[注]弗洛伊德对“我们清楚自身的理由”这一观念的质疑得到了有力的支撑。

美国心理学家理查德·尼斯贝特（Richard Nisbett）和蒂莫西·威尔逊（Timothy Wilson）1977年发表了一篇非常有影响力的文章，文章中回顾了大量的证据，这些证据表明我们对自身心理过程只有极少，甚至压根儿就没自省，还表明我们对这些过程的口头表述通常是虚构的。[注]实际上，他们认为，我们解释自身行为的方式与解释他人行为的方式没有太大不同。在解释他人行为时，我们考虑了对他人及情境的了解，接着寻找貌似合理的原因（受到我们文化中认可的因果解释类型的影响）。在解读自身心智、解释自身的行为时，我们亦是如此（利用更丰富但无根本差异的证据）。哲学家彼得·卡拉瑟斯（Peter Carruthers）在其著作《心灵不透明》（*The Opacity of Mind*）[注]中，展示了最近的研究究竟对尼斯贝特和威尔逊的理论（他从实证和哲学的视角对该理论进行进一步说明）证实并强化了多少。

我们对他人的推论往往非常有洞见，而对自身的推论则再糟糕不过了。我们通常可以成功找到少量的、的确对自身的信念及决定发挥过作用的信息。我们犯的系统性错误是假定能直接自省心理状态及其产生过程。

不能自省的、普遍无意识的过程的存在对我们自身的常识观念形成了多大的质疑呢？“意识无法触及知觉、记忆或运动控制”这一既定事实真的不是问题所在。令人更加不安的是，我们发现，即使在看似有意识的选择中，我们真正的动机可能也是无意识甚至无法被自省的；在许多情况下，我们充分相信的理由不过是事后的合理化而已。

我们（在第二章中）已在推理心理学中见过这种事后合理化的明显例证。在沃森四卡选择任务中，参与者在开始推理前凭直觉去选择卡片。即使这一选择任务的所有版本在逻辑上完全一样，参与者的选择却在有些版本中是对的，在有些版本中是错的。当被要求解释其选择时，他们轻易就给出了理由。当其选择碰巧正确时，他们给出的理由便是合理的，反之则是编造的。在这两种情况下，合理的和编造的理由显然都是事后的合理化。而且，参与者均没有意识到，实验证据显示的这些因素实际上促使他们做出了最初的选择（且不论他们的回答正确与否，因素是相同的）。当然，即使这些实验数据非常强劲有力，但仍带有实验室的味道。

幸好，并非所有的实证研究都与现实生活无关。1964年3月13日，基蒂·吉诺维斯（Kitty Genovese）在纽约遭到残忍杀害，当时许多邻居听见了她的呼救声，却没有人管，这就促使社会心理学家约翰·达利（John Darley）和比伯·拉塔奈（Bibb Latané）开始研究，在何种条件下人们会或不会施以援手。[注]他们发现，当有许多人可以帮忙时，实际上每个人去帮忙的可能性就会降低。在场的旁观者促使人们对当事人的痛苦置之不理（达利和拉塔奈称为“旁观者效应”现象），但是这一起因人们通常并未意识到。

在拉塔奈和朱迪斯·罗丁（Judith Rodin）开展的一项研究中，[注]参与者被告知会去参加一个有关游戏的市场调研。和善的助手在实验室门口单独接待他们，并将他们带到与自己的办公室相连的房间，让他们填写调查问卷，之后便回到自己的办公室，在那儿参与者能听到翻页的声音、开抽屉的声音等。过了一会儿，参与者听到该助手爬上椅子，又听到一声巨响，紧接着是一阵叫喊：“啊，上帝呀，我的脚……我……我……移不动它。啊……我的脚踝！……我……不能把这个……东西……移开。”

一种情况是，当所有这一切发生时，参与者独自一人在房间里。另

一种情况是，有一位男士跟参与者同处一个房间，表现得也像一个参与者（而实际上是实验者安排的“托儿”）。这位男子几乎不理会这声巨响和叫喊声，只是耸耸肩，然后继续做问卷。当真正的参与者独自一人在房间时，70%的人都会去帮忙。但当他们与这个明显冷酷无情的参与者共处一室时，只有7%的人会去帮忙。

实验结束后，参与者很快就他们的反应接受了采访。大多数出手相助的人回答说：“我不能十分确定到底发生了什么，但我想我至少应该去弄清楚。”大多数未出手相助的人称，他们认为不管发生了什么事，肯定不是太严重，况且若有需要，同在隔壁办公室的人会帮忙的。他们没有因此感到特别内疚或者不自在。当然，如果这真的是一个紧急事件的话，他们是会帮忙的（他们就是这样声称的）。

当被问及是否房间里有另外一个人会影响他们做出不去帮忙的决定时，大多数人都坚称根本没有影响。然而，我们知道，相比有他人在场时，他们独自在房间，有更多的人出手相助的可能性要大10倍以上。换句话说，他人的存在对他们的决定有巨大的影响。许多因素有助于解释这一“旁观者效应”：当有其他人在场帮忙时，大家的责任感就会变淡，其他人不做出反应可能会使那个出手相助的人显得可笑等。这一事例强有力地证明，在界定促使人们挺身而出或漠然置之的因素上，人们错得有多离谱（在第十四章可以看到更多惊人的例子）。

发生了什么？我们就是因为一些通常无意识、不能自省、重构时犯错风险较高的心理学理由而形成信念并做出决定的吗？或者说，通常发生的情况与反映理由在我们心智活动中作用的常识性观念分歧更大？

模块没有理由

到目前为止，上述所有证据表明，人类对指导他们的理由知之甚

少，而且经常弄错这些理由。我们想要对常识图景发起更彻底的挑战。我们不是错在经常错判真正的理由，而是错在一开始便假定所有的推论都受到理由的指导。我们想要表明，理由不是在直觉推论过程本身而是在对直觉的事后解释和辩护中发挥重要作用。

当然，几乎没有哲学家或心理学家会否认一个明显事实：我们通常在形成观念及做出决定时并没有意识到这样做的理由。尽管如此，他们仍会主张，不论我们是否意识到自身理由，仍然会受到理由的指导。只是这些理由是“隐性的”而已。这就是在直觉推论中发生的事情。但什么是“隐性的理由”？它们在没有被有意识地表征时，如何能发挥所谓指导作用？

“隐性的”这个词是从语言沟通研究中借用的，这类研究对该词有相对明确的意义界定。但是，当心理学家或哲学家谈起“隐性的理由”时，他们的意思要么是这些理由被无意识地表征，要么是根本就没有被表征（虽然不知为何仍有关联）。这一模棱两可的情况通常有待解决，而且谈及“隐性的理由”，不过是认可常识性观念的一种方式而已。常识性观念认为，人们的想法和行动一定在某些方面是基于理由的，而对这种理由的心理现实与实际作用并无任何确切观念。[注]

我们认为“显性的”和“隐性的”区别只有在语言沟通研究中是明显且有用的。我们还认为，人们所谈的“隐性的理由”的唯一明显有用之处是，理由被含蓄地传达。譬如，当智恩回绝罗博给予奶昔的好意时，说道：“谢谢，但是你知道，我们韩国人大多数都受不了乳糖。”她拒绝的理由就是隐性的，用了“隐性的”语言学意义：这一回答不是直截了当的，因为她并没有说自己受不了乳糖，但是可以而且也应该能从她的言语中推断出来。

尽管如此，人们仍会坚持认为，心理学理由可以是有意识的，也可以是无意识的，认为至少某些无意识的理由可以被当成有意识的，并认

为把无意识的理由当作有意识的“隐性的理由”也说得通。但是从这个意义上来讲，真的有隐性的理由吗？我们表示怀疑。

我们认为无意识的、直觉的推论由专门的模块来执行。模块输入的是特定事实的表征，并利用专门的程序从这些表征中得出结论。当模块运转良好并产生合理的推论时，在模块输入中所表征的事实，的确客观上会支撑模块产生的结论。然而，这一观点却与某主张大相径庭，该主张认为特定事实的表征是指导模块运作的无意识理由。

下面说一下为什么。某事实（任何事实）不只是支撑某一结论的客观理由，还是支撑各种不同结论的客观理由。譬如，“今天是星期五”这个事实是一个客观理由，不仅能从中推断出“明天是星期六”，而且还能推断出“后天是星期天”；可以推断出今夜犹太教的安息日（Jewish Shabbat）要开始了；还可以推断出许多办公室职员今天的穿着相比本周的其他工作日更随意；以此类推，以至无穷。

而且，同一事实可能是支撑某一结论强有力的客观理由，但支撑另一个结论时很无力。例如，“梅子成熟”这个事实虽是推断出“如果梅子不被采摘便很快就会掉落”的强有力理由，却也是推断出“这些梅子马上就会被采摘”的较为无力的理由。同一事实甚至可能是支撑两个矛盾性结论的客观理由。例如，“在下雪”这个事实可能是“待在家”的理由，也可能是“去滑雪”的理由。

由此可见，单纯事实的心理表征并不能单独构成心理学理由。只有当事实被表征为支撑某一特定结论时，这一事实的表征才是心理学理由。你知道正在下雪，但不知道对你而言，这个事实是待在家、去滑雪、做其他事的理由，还是只是单纯的事实，且根本不是任何特定结论的理由支撑。我们不能在不知道他人理由的具体功用时就贸然归因。好吧，你可能会想，那又怎样？当然，如果某一（真实的或假定的）事实的表征作为得出某结论的前提，那么它恰恰就是支撑这一结论的心理学理由。然而，这一说法却是错误的。信念作为得出某一结论的前提，却

不一定是支撑这一结论的心理学理由。

伊本·艾尔-海什木很久以前便主张，大脑通过执行三段论的步骤来执行无意识推论。如果真是这样的话，那么特定事实的表征会作为该三段论的小前提，能从这一特定事实中推断出结论的规则会作为大前提，这两个前提一起，可以视为支撑该三段论结论的心理学理由（每个前提只是部分理由）。

对所有推论（有意识的或无意识的）的这一逻辑学上的解读仍然得到普遍认同。也许这便是无论何种信息作为推论的输入信息（或“前提”）一定是支撑这一推论结论的理由这一点看似不言自明的原因。

可是，比较心理学及认知心理学的现代研究已彻底推翻了认为“无意识推论的产生是通过执行三段论步骤，或者更普遍来说是通过逻辑推导步骤”的观点（正如我们在第五章中主张的那样）。

我们认为无意识推论产生于模块，这些模块利用存在于作为输入的特定事实与作为输出的结论之间的惯常联系，但是它们既没有把这一联系表征为三段论的大前提，也没有表征为三段论的其他任何形式。我们指出，特定事实的表征不能单独作为支撑任何特定结论的心理学理由。无论怎样，模块都不需要理由来指导。它们能利用事实的表征作为输入，既不以表征为理由，也不以表征为这些事实与其推导结论之间的关联。模块不需要激励或指导来大量炮制其输出。

我们先来以初步推论的案例为第一个例证。即便初步推论发生在神经系统中，确切地说，也不会发生在心智中。排汗是人类身上（和马身上）发育良好的机制。(注)当体温过高时，下丘脑（某脑结构）引发排汗功能来给身体降温。很显然，这事确已在我们身上发生，而非我们欲做之事。尽管如此，排汗机制完成了初步的有效推论：它输入各种神经探测器提供的体温信息，当时机合适时，便输出对汗腺的指示。在此过程中，排汗机制利用了“在一定的温度临界值之上，流汗是合宜的”这个普

遍事实，但它并不能表征这一普遍事实。表征在模块中的当前体温信息不能单独作为任何特定之物的理由支撑。很明显，模块的运作不受任何心理学理由的指导，且不需要任何此类理由来发挥作用。

我们选取的第二个例子是沙漠行军蚁。沙漠行军蚁一旦找到些许食物，便会以近乎直线的形式加速返回巢穴（正如我们在第三章中见到的那样）。沙漠行军蚁此举多亏了维纳说的“导航工具箱”，即拥有专门子模块的复杂认知模块。子模块使用的每个步骤（如计算步数、将轨迹的角度变化考虑进去等）都已进化，得以利用可靠的规律，却不用表征这一规律。沙漠行军蚁在巢穴外的每一次突袭中，这些子模块推断出的相关信息就会自动融合，并帮助沙漠行军蚁确定返回路线。在解释这一进程时，并无理由假定沙漠行军蚁受到显性理由的指导或所涉模块受到隐性理由的指导。

你可能会问，为何我们要费事地明确指出，人体排汗和沙漠行军蚁定位没有受到理由的指导，不论是显性的理由还是隐性的理由呢？因为与这两个案例相关的考量很自然就拓展到了不那么直接的案例中，而且从一开始就拓展到了认知推论的案例中。

还记得我们在第一章中谈到的阿德尔森棋盘错觉（参见图1.3）吗？参与者被问及棋盘图片中的两个正方形哪个更暗，而实际上，二者的灰度恰恰是一样的。他们认为两个正方形中的某一个比另一个更亮些，这便是错觉。他们之所以这样认为，是因为他们不仅根据棋盘反射到他们眼中的光线，还根据他们假定棋盘所接收的光线，来推断棋盘表面的相对暗度。两个正方形中的其中一个画在圆柱体的阴影部分，因此接收到的光线比另一个要少。参与者自动补充了在光线接收上的这一差异，然后认定此正方形比实际的要亮。

在这个关于认知推论作用的教科书式范例中，没有人会说我们是用有意识的理由来认定一个正方形比另一个正方形暗的。依照正方形反射的光与其接收到的光的比率，相关模块算出了每个正方形的相对暗度，

因为该模块的进化使然；在没有错觉的案例中，该模块能这样做是存在客观理由的；不过，这些理由并没有在该机制中被表征。该模块并未受到无意识的心理学理由的指导。

我们希望读者能认可，某些基本的认知推论并不受到心理学理由的指导。可是，总体而言，这一论断会拓展到认知自动化程度更弱、更具诠释性的方面，并拓展到直觉推论吗？当拉塔奈和罗丁实验中的参与者听到来自隔壁房间的不同噪声时，他们辨别出其中一些噪声是撞击声，另一些是说英语的声音。可以假定，这个辨识是完全无意识的：心理表征的理由在这一过程中不起任何作用（即便它是一个推论过程）。将这些感知整合到一起并推断知悉隔壁的女士摔倒了且受伤了，需要更多阐释，知悉她可能需要帮助所需的阐释比前者只多不少。这是否意味着参与者要得出这些更深层次的阐释性结论必须要有显性的理由，或意味着所涉机制必须要有无意识的理由？

任何熟悉这种现代实验环境并具备社交能力的人，都会发挥他们的模块分析能力来做出相关的推论。他们会自发辨别出这些声响不仅是撞击声，而且若结合微妙的声学特点来看，还是硬家具与较柔软的身体撞击的声音，如人体落在坚硬表面上。我们认为，允许这种辨识的认知推论程序利用了噪声特点与其可能原因之间的关联，且并未表征这些关联，更不用说将其作为理由来表征了。

当参与者听到“啊，上帝呀，我的脚……我……我……移不动它。啊……我的脚踝！……我……不能把这个……东西……移开”这些话时，他们会自发地把“它”理解为说话人的脚，把“这个……东西”理解为某个巨大且坚固的物体，如落在说话人脚上的某件办公室家具，这样理解就远远超出了“它”和“这个东西”的语言编码含义。为执行这些语用推论，参与者用到了利用言语沟通中的关联性规律且并未表征这一规律的模块理解程序。[注]语言性阐释众多，他们只专注于其中一种，并未有意识或无意识地表征这样做的理由。

运用这一直觉方式，在了解到隔壁那位女士受伤了且正在表达自身的窘迫之境后，大多数参与者会自发地推断出该女士需要帮助。我们再次表明，他们不用心理表征（有意识地或无意识地）支撑推论的理由，便能得出结论。事实上，如果某个人必须心理表征一个这样的理由，才能意识到一个处于痛苦和抱怨中的人可能需要帮助，那么这种无意识的同理心的缺失，可能会被看作认知能力和社会能力受损的症状。

尽管如此，毫无疑问，在拉塔奈和罗丁的实验中，理由的确发挥了作用。问题是，在这些事件进程中，理由何时才会出现。听到隔壁办公室的撞击声后，一些参与者站起来去看看能否帮上忙。其他人仍然坐着。他们的决定（不论是尝试去帮忙还是什么都不做）都是无意识的决定（即在做出决定时没有考虑到理由）吗？我们认为，除非他们表现出一丝犹豫，否则便没有充足的理由认定他们受到表征在心智中的有意识或无意识理由的指导。虽然可能几分钟过后，当实验者问他们当时怎么想的时候，他们很乐意提供那些他们认为的激励其行动或不行动的理由。

当一些参与者（特别是那些出手相助的）被问及时，他们可能已经想出了一些理由。对他们而言，找到理由是毫不费力的：在某些情况下，很明显会有人需要你的帮助，“想要去帮忙”是一个明显充分的理由。那些坐着不动、什么都没做的人，需要花时间来思考可以为自身辩护的理由。他们可能已感觉到很有必要去想一个正当的理由，尽管对他们而言，想出一个貌似可信的正当理由并不是那么难。在这两种情况下，通常都是在做出决定后才来思索支撑其决定的个人理由。如果如我们所料，理由的作用不是去指导形成信念及做出决定，那么理由的作用何在？

理由是为了社会消费

人们的所作所为影响着他人对他们的评价，换句话说，影响着他们的声誉。这些对声誉的间接影响可能会和他们直接行动目标同等重要，无论这个行动目标是什么。有社交能力的人非常关心他们的行为将被如何解读。通过解释及为自己辩护，人们会捍卫甚至提升自己的声誉。如果没能这样做，他们的名誉就有可能会受损。

人们经常主动为自己的行为想出充分的理由，期望被叫去为自己解释或辩护。一旦你参与到可能伴随着声誉代价的行动进程中，一个不同的心理机制可能就开始运作了，有时甚至在这一进程之前，当你尚在筹划时便已如此。该机制的功能是助你荣获声誉，从而提供能为你的行为辩护的解释。在拉塔奈和罗丁的实验中，无视隔壁房间动静的参与者，可能已经意识到他们被动的反应可能会招致批评。这就会促使他们找理由来为自己辩护。

声誉管理机制表现得就像不论你做了什么都会为你辩护的律师。尽管如此，一有机会，律师便可能建议长期委托人放弃很难或不可能予以辩护的行动方案。当声誉机制不能给出充分的解释时会怎样？已经想好的（且可能已同意的）行动方案会产生最初在决策过程中并未考虑到的代价。在这种情况下，若声誉机制没能拿出充分的解释就可能会产生反馈效应，导致最初的决定被取消，至少是调整。譬如，在拉塔奈和罗丁的实验中，可能少数参与者最初没有考虑帮助隔壁房间陷于痛苦中的女士，紧接着便意识到，这可能会很难为自己的消极反应做辩护；他们犹豫不决，权衡着两种反应的辩护难度；最终还是做出了选择。

一些有趣的实验证据表明，搜寻旨在辩护的理由实际上会影响行动（详情请看第十四章）。尽管如此，你想做出的自我描述与符合事实的描述并非完全一样。

所以，当人们拿出理由来解释及辩护其信念或行动时，给出的解释可能会与其心智中发生的真实情况相冲突，这一冲突表现为三个不同的方面，每个方面都有各自的有趣之处。

第一，人们通常认为，他们在形成信念或做出决定的过程中就已经考虑到了理由并受其指导，而实际上，这一信念或决定是基于直觉形成的。我们所有人在此犯的错误是，错误地假定我们在做出直觉推论时直接获知了心智的真实想法，错误地假定我们意识到了指导我们的理由，错误地假定我们擅长自省。与这样一些假定相反，在形成直觉决定或信念时，我们并没有受到理由（甚至包括“隐性的”理由）的指导。

即便如此，我们的确经常正确地识别推论过程中作为输入的信息。当遇到这种情况时，我们犯的错误只是将这一信息描述为个人理由。那些去看能否帮到隔壁房间女士的参与者，在心理上表征了“那位女士处于痛苦中且需要帮助”这一事实；他们当时并未在心理上表征的是高阶考量，即“那位女士需要帮助”是去帮忙的理由。“帮助有需要的人”的正当性（至少在某些情况下）显而易见，因此在社交能手的身上，这一点被不用表征便能运作的专用程序所利用。对于我们的社会互动而言，这种错误的理由自我归因根本不重要。实际上，与其说是个人的失误，不如说是社交鼓励下的常识心理学。但是，从科学的观点来看，这应当被视为一种误读：自发地去帮助他人不需要受到有意识或无意识理由的指导，只是投桃报李而已。

第二，人们可能弄错了推论中作为输入的信息。那些出手相助的参与者主要是受到了房间中另一个在场且无动于衷的人的影响，但他们没能承认这一至关重要的因素。很可能当他们声称其决定受到了“没什么太重要的事情发生，总之不需要我们的帮助”这一想法的激励时，只是在报告他们心中的想法，而这些想法出现在他们做出“使自己的行为与房间中的另一个人保持一致”的直觉决定之后。在向实验者说出这些理由时，他们可能是诚实的，即便有点自欺欺人：那些自以为友善并乐于助人的人和那些意识到本应该出手相助却未采取行动的人，可能会为自身行为感到困惑。解决这一困惑非常简单的方法便是，假定他们当时一定有看似充分的理由来认定不需要他们的帮助。

第三，存在没能为自己欲做之事找到足够合理的理由支撑的人，结果这些人变得摇摆不定，进而改变或至少调整了自己的行动方案。在这些情况下，人们援引的个人理由真的在他们的最终决定中发挥了因果作用。这些理由没有直接影响最终决定，却阻碍了最初决定的执行。尽管如此，与这些人的信念和声称不同的是，真正的因果过程不是从理由到决定，而是从试探性的决定到正当理由的搜寻，再到搜寻失败，然后调整决定，旨在形成更容易辩护的行动方案。

到目前为止，我们所言都是为了暗示，在人类认知中，无意识的理由没有什么作用。如果无意识的理由没有什么作用，那么它很有可能不存在。如果无意识理由不存在，那么隐性的理由也就不存在。当我们把隐性的理由归于自身时，只是在依据新构造的有意识理由来解读自身想法及行为，我们把这些理由虚构性地表征为先前就是隐性的。

难道这就意味着心理学理由，不论是“隐性的”还是有意识的，都不是真实的？都只是纯粹的构造？不，理由的确是构造出来的，却有两个限制性因素，能够确保理由包含某种程度的心理现实和社会现实。我们援引的理由必须在心理学上有意义。谈及理由时不需要（事实上，我们认为是不能）对我们心智中真实发生了什么给出准确的解释，但理由通常会突出那些发挥因果作用的因素。理由通常由少量心理学洞见构造。

而且，为实现其辩护功能，心理学理由必须把社会公认的客观理由表征为充分的理由。考虑到社会因素和文化因素，这个社会认可可能是有失公允的，但如果是纯粹武断的事情，理由是如何发挥作用的呢？对许多信念和惯例而言（特别是对于用理由来辩护或作为论据的信念和惯例），我们坚持认为真正的客观性是文化成功的一个重要因素（这一事实我们会在解释理由的评价作用时进行很好的阐释，在第十三章中会讨论）。

把理由归于某人（自己或他人），与其说是形成对事情形态的假设，不如说是为社会行动构造一个工具，用于辩护、评估或论证（我们

将在第八章至第十章谈到论证）。但是，是什么使理由成为有用的工具的呢？

当我们给出支撑自己行动的理由时，不仅是在为自己辩护，也是在让自己承担义务。首先，通过援引理由，我们为自己描述的想法和行动（就如同我们有理由采取的态度和行为）承担起了个人责任。由此我们表明，希望他人要么认可我们有权想我们所想、做我们所做，要么准备好质疑我们的理由。当所想或所做不可能得到认可时，我们便通过给出理由亮明了自己的防线：我们有理由，即便不充分，但至少当时看似充分。基于理由的防卫通常让我们在负起责任的同时否认罪责。

通过给出理由，我们也对自己未来的想法和举止做出承诺。我们援引理由来作为过往想法和行动的动机，表明认可这些理由的规范（标准）适用性，并承诺未来受到相似理由的指导。对受众而言，我们承诺负起责任，并在未来受到所援引的、解释过往的理由的指导，这比起我们所谓自省的准确性意义更加重大。这就解释了为何我们大家都关注他人的理由，为何我们会拿出自己的理由。

用更加社会性的术语来说：理由是社会的构造物。我们歪曲、简化对心理状态及其因果作用的理解，并注入强剂量的规范性，以此来构造理由。调用和评估理由有助于协商并记录个人的想法、行动、责任和承诺。这种社会记录涉及谁在想什么和谁做过什么，它既包含共识也包含异议，但在指导合作性或对抗性互动、影响声誉及稳定社会规范中发挥着主要作用。因此，理由主要是为了社会消费。

---

1. 单词“reason”（理性、理由、原因）广义上来说也可以当作“cause”的同义词（理由、起因），就像在句子“今夜天很黑，是因为今天是新月日”中一样。在本书中，当会造成混淆时，我们都不会使用“reason”的广义概念。
2. 欲了解最近研究理性且与我们的研究相关的心理学著作，可参见：Alvarez 2016。
3. Nagel 1976; Pritchard 2005; Williams 1981.

4. 例如，可参见：Baars 1993; Block 2007; Carruthers 2005; Chalmers 1996; Dehaene 2014; Dennett 1991; Graziano 2013; Rosenthal 2005。
5. Nisbett & Wilson 1977;也可参见：Greenwald & Banaji 1995。
6. Carruthers 2011.
7. Darley & Latané 1968.
8. Latané & Rodin 1969.
9. 汉斯·拉克兹（Hannes Rakoczy，2012）突出了显性或隐性特性的各种用处引起的困惑。佐尔坦·狄恩斯（Zoltan Dienes）和约瑟夫·珀纳（1999）用了精准但又有点另类的方式来定义“隐性的”，这在下一个语境中需要进行详细的讨论。
10. 可参见：Jessen 2001。
11. Sperber & Wilson 2002.

第八章
理性能否成为一个模块

这天，阳光明媚，来自南埃塞俄比亚多尔兹部落的一位长者在劝诫一群无所事事的年轻人：“你们不应该抽烟，这违背你们的宗教信仰！”

这群年轻人便指出，年轻的法国人类学家丹（Dan）过去几个月一直住在多尔兹部落，此刻他就坐在附近，在做什么呢？抽烟！长者随即来到这个外国人身边，要求得到一个解释：“为什么你要抽烟？”这个人类学家不知道该说什么。“好吧，”他含糊地说道，“我父亲抽烟，我的祖父也抽烟……”

长者又欢欣鼓舞地回到那群年轻人身边：“你们看到了，这个外国人是出于对他父亲和祖父的尊敬。他们抽烟，所以他才抽的。遵循这一惯例，出于对你们祖先的尊敬：他们从来不抽烟，所以你们也不能！”

正如这个例子表明的，同样的理由能用来为自己辩护，也能作为论据说服他人。（这里特别需要指出的就是，当与其他的考量相结合时，同样的理由可以用来为个体行为做辩护，也可以据此反对其他有此行为的人：理由的重要性视具体语境而定。）

回顾性和预期性地运用理由

有时，人们用理由来解释或辩护已采取的决定及已持有的信念。这就是回顾性地运用理由。有时，人们用理由作为论据来支撑新的决定或

新的信念。这就是预期性地运用理由。当理由被预期性地运用时，可能便是要回答某人自己至今仍不知道答案的问题，或者说服他人相信自己持有的观点。其中，第一种情况是探究式推理，第二种情况是论证式推理。图8.1便展示了理由的这些运用。

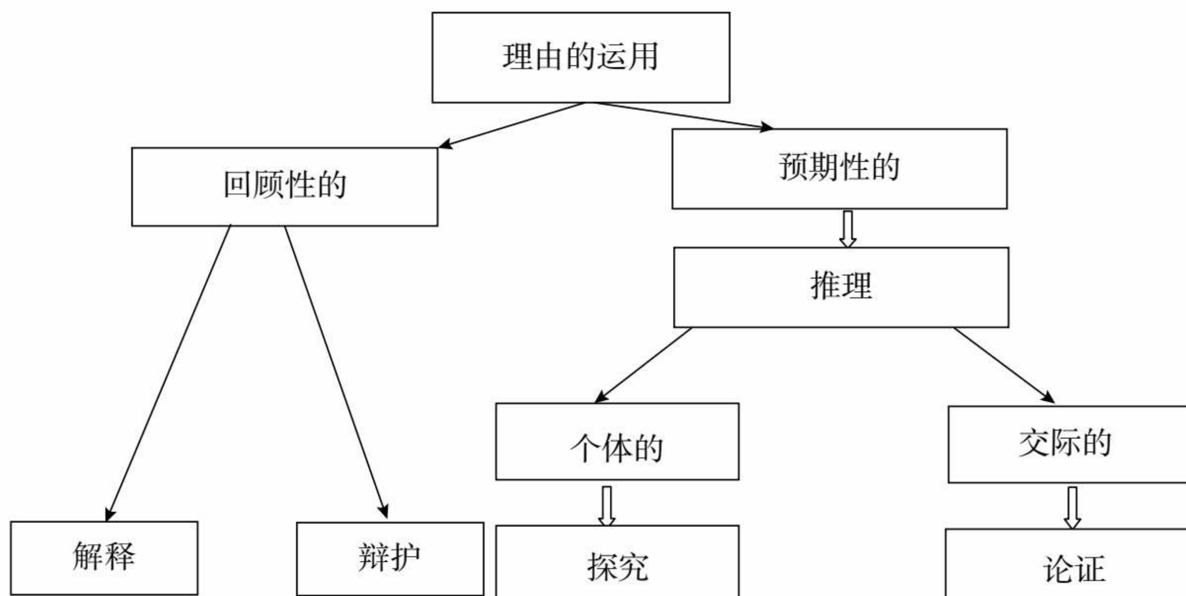


图8.1 回顾性和预期性地运用理由

哲学和心理学文献把理由的回顾性和预期性的运用看作两个完全不同的话题，很少注意到或压根儿没有注意到二者的共同之处。与之相反，我们认为，理由的回顾性运用和预期性运用没有分界线，且有许多重叠的情况。人们通常认为理由的解释作用比辩护作用更加重要（或至少同样重要）。与之相反，我们认为，理由的辩护作用比解释作用更重要。人们认为，旨在独立回答问题的个人探究式推理是主要的，而旨在交际的论证式推理是次要的。与之相反，我们认为，论证是主要的。

推理本身便是哲学及心理学中的主要话题。主流观点认为，推理旨在获取真理并做出好的决定，而且其不偏不倚、公正客观。运用在推理中的理由应该是非个人的“论据”，这些论据的效力在于自身的形式特性（这在逻辑和概率论中都研究过）。但是，运用理由来解释、辩护已采纳的观点或已做出的决定，及据此采取的行动，会受到个人因素的影

响，并不能真的做到不偏不倚。

非个人的推理与个人辩护之间的对照有何意义？个人考量在有效的推理中总是不相干吗？在实际推理中，至少个人考量的相关性通常得到了认可。研究实践理性和道德哲学的哲学家，特别是约瑟夫·拉兹（Joseph Raz）㊟，更愿意采用统一的方式来对待理由的回顾性运用与预期性运用，这一点不足为奇。

总之，用统一的框架来处理回顾性和预期性的理由，这一点有简单的论据做支撑。这一论据是：所有可以用来辩护的理由都可以用在推理中，反之亦然。如果你心存疑虑，那么就设想一个能够很好地为“大家都不应该抽烟”这个观点做回顾性辩护的理由。当然，这一理由之后可以被预期性地用来说服自己或他人相信“大家都不应该抽烟”。相反，想象一下，某位数学老师问其学生是否存在最大的素数，然后帮助他们推理并重新找到不存在这样一个素数的证据。无疑，使学生一开始就得出“没有最大的素数”这一结论的同一证据，此后会变成支撑其信念的无可指摘的正当理由，就如已经在老师身上实践过的那样。

也就是说，在标准的论证案例中，在拿出理由说服他人时，同样的理由既与回顾性相关，又与预期性相关。论证者是在试图说服受众接受他已持有的观点。譬如，奥拉夫（Olav）和利维亚（Livia）打算点一盘生蚝，奥拉夫就在纠结他们应该点哪种红酒来搭配。他问利维亚，利维亚回答：“密斯卡岱（Muscadet）白葡萄酒！这种酒的酸度和矿物风味与生蚝搭配非常好。”奥拉夫曾听说关于密斯卡岱白葡萄酒不好的言论，但是鉴于利维亚在红酒方面比他更在行，因此利维亚的论据说服他相信应该点密斯卡岱白葡萄酒，此外他还应该修正自己的负面看法。利维亚说服奥拉夫的论据也是为她个人观点辩护的理由。真正的论证者使用的说服受众的论据是她认为能为自己的观点提供充分回顾性辩护的理由。

我们想要证实，回顾性的辩护和预期性的推理不仅在很多方面重叠，不仅利用同样的理由，而且依赖同样的机制、同一个模块来传递对理由的直觉。

理由本身必须被推断出来

“除了人类，动物也思考理由”这一发现会令人十分惊讶（且饶有趣味）。我们已经表明，理由在人类思维中占据重要位置，是因为其在人类十分丰富且复杂的社会互动中发挥着独一无二的作用。理由有助于形成个人责任、相互期待和规范。然而，说这些并不是要表明人们如何能知道他们的理由（即便正如我们看到的那样，这是一个非常不完整的认识）。

正如我们在第七章中指出的那样，比起仅认可某一事实，找到理由需要付出更多的精力。你能走出去，而后看到路面湿了，但你不能简单地把这当作下过雨的理由。“路面湿了”可以作为支撑“路面很滑”“室外温度在冰点以上”“人们的鞋子会弄脏”等推断的客观理由。想一想，“路面湿了”不足以构成一个独立的理由。此外，你可以从“路面湿了”这个事实凭直觉推断出下过雨，而不用在这一过程中在心理上表征前提和结论间的联系。只要你抱有“从路面湿了这个事实可以推断出一定下过雨”的想法，你就会识别支撑你结论的理由。

假设你真有支撑“下过雨”这一推断的理由，问题仍然存在：你如何得知“路面湿了”这个事实是支撑你结论的理由？理由不是借助魔法出现在我们的脑海中的。只有通过别的什么，比如另一个高阶推论，才能识别某一事实是推断出给定结论的理由。

所以，理由是如何被推断出来的？是通过找到更深层次的理由来支撑我们的理由吗？有时当然可以，但大多数时候不能。假定理由的识别

本身必须总是基于会导致无限倒推的高阶理由：要推断结论A，需要理由B；要推断“B是推断A的理由”，需要理由C；要推断“C是推断‘B是推断A的理由’的理由”，需要理由D；以此类推，以至无穷。因此，理由的识别最终必须基于直觉推论而非进一步的推理。这个无限倒推的观点之前就有人提出过：刘易斯·卡罗尔（Lewis Carroll）［《爱丽丝梦游仙境》（*Alice in Wonderland*）的作者］于1895年发表过一篇短小精悍的注解，题为“乌龟对阿基里斯说的话”（What the Tortoise Said to Achilles），注其中包括该论点的早期版本。但有个问题从未被强调过：所有这些有何心理学暗示（若有的话）？

我们认为世界是有规律的（像物理学定律这种普遍规律、巴甫洛夫条件反射实验中铃铛—食物关系这种局部规律），因此推论便成为可能。在推论过程中，使推论得以进行的规律不必表征为推论过程中的前提，而是蕴含在专门的程序中。直觉推论产生于利用这种程序的自主模块。从这一视角来看，“理由的识别基于直觉推论”这一事实表明，一定有模块能用来识别理由的规律。真有这样的模块吗？如果有，理由的规律是什么？有更好的替代性解释来说明理由是如何被识别的吗？在推理心理学中，这些问题甚至从未被讨论过。

心理学家已研究了人们表征前提的范式及从这些前提中推断出结论的方法（这就是我们在第一章中说的“心智建模师”和“心智逻辑学家”间的辩论）。很少有人研究，人们何时及如何推断出“特定前提为特定结论提供理由支撑”。“理由最终必须基于直觉推论”这一事实遭到忽视或被视为无关紧要，仿佛理由的这个最终的直觉根据严重脱离了实际推理过程，所以对心理学研究没有意义。实际上，事实恰恰相反。

在日常推理中，对理由的高阶推理实属罕见。人们运用的大多数理由都是直接基于直觉推论。譬如，“路面湿了”是推断出（有犯错的危险）“下过雨”的理由，这在直觉上就显而易见。当理由的直觉根据是间接的时，关联事件就会较少。一个吹毛求疵的推理者会额外加上一步：

从直觉上看，对“路面湿了”最可能的解释是“下过雨”，因此“路面湿了”是推断“下过雨”的理由。即便在人们确定会推出理由的更正式推理中，直觉根据也从未走远。人们凭直觉识别理由的能力使推理不论在原则上还是在实践中都是可行的。

我们在第二章中细究了埃文斯、卡尼曼和斯坦诺维奇等人的双重过程理论，该理论的核心是，假定直觉推论和推理是通过两种截然不同的机制来实现的。对于这一点，我们并不认同。本书的主要主张之一是，推理不是直觉推论的替代品，而是对理由的直觉推论的运用。

我们主张，使人类能推断出自身理由的是其元表征性直觉推论能力。为清楚地表达这一主张，我们再次讨论并拓展了第三章至第五章中介绍的观点，这一观点有关三个对理解人类心智必不可少的话题：直觉、模块和元表征。

对理由的直觉

我们在第三章中已经表明，直觉既不产生于一般直觉官能，也不产生于特定类型的推论过程，而是各种推论模块的产物，其产物在某种程度上是有意识的，尽管其运作仍然是无意识的。我们现在面临的问题是：存在这样一个能对理由做出直觉推论的模块吗？为了回答这个问题，我们首先必须强化对直觉的总体认识。

产生直觉的推论机制，即通过无意识的过程得出有意识的结论的机制是多种多样的。直觉推论没有共有的本质特征，其他类型的推论却有。下面将举出两个推论案例，很明显两个都是直觉的，但除此之外，二者鲜有相同之处。

神经系统科学家维拉亚纳·拉马钱德兰（Vilayanur Ramachand-ran）和心理学家爱德华·哈伯德（Edward Hubbard）展示了沃尔夫冈·柯勒

［Wolfgang Köhler，格式塔（Gestalt）心理学派的奠基人之一］在早期的著作中画的两个图形（见图8.2）：

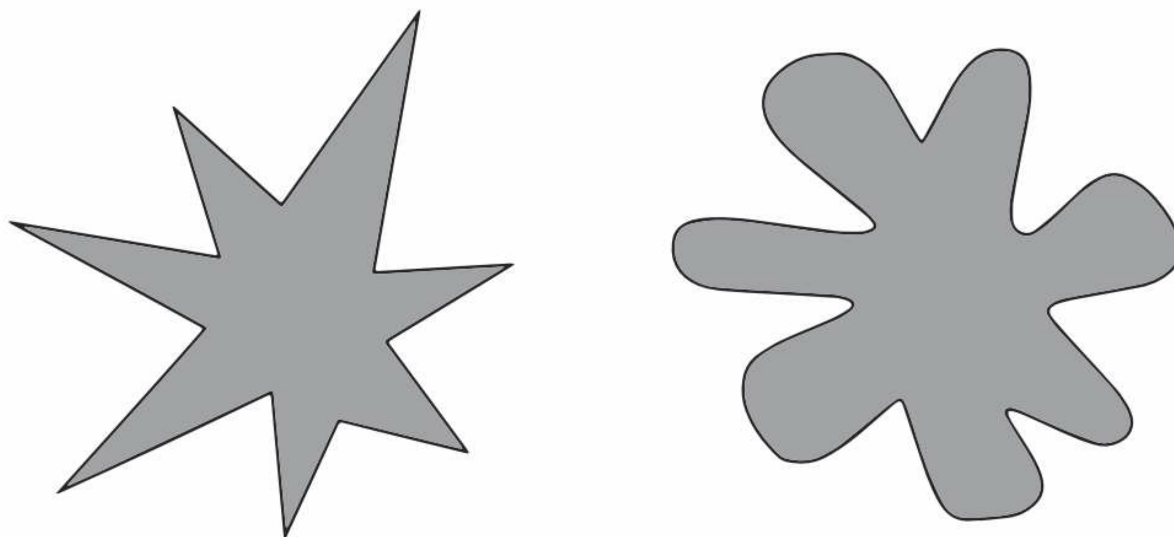


图8.2 Kiki和Bouba

他们对人们说：“在火星文中，这两个图形一个是Bouba，另一个是Kiki。猜一猜它们分别对应哪个图形。”95%的人认为左边的图形是Kiki，右边的图形是Bouba。(注)做出这种判断的强烈直觉似乎是基于声音和图形的联觉。在此，直觉与知觉相近。

Bouba-Kiki直觉十分具体。其他的直觉都很抽象。英国哲学家乔治·爱德华·摩尔（George Edward Moore）曾于1942年指出，做出“P，但我认为不是P”这种论述是荒谬的，比如“今天是星期一，但我认为不是星期一。”这一观察结果在直觉上很明显，但却没那么容易解释。与之看起来相反，这一命题不是自我矛盾的，“今天是星期一，但我认为不是”这一论述极有可能是正确的。荒谬之处并不在于所陈述的事情，而在于陈述时采用了第一人称和现在时态。这样就清楚多了，但还不足以解释荒谬的直觉。实际上，虽然对“摩尔悖论”的直觉是没有争议的，但时至今日，其解释仍然是哲学家争议的话题。(注)

正如Bouba-Kiki和摩尔悖论这两个例子阐明的那样，促使一些结论

能够被直觉接受的，既不是这些结论的内容，也不是其形成方式，而是我们对这些结论的信心。直觉的区分不是基于认知特点，而是通过元认知特点。注

我们对直觉的自信不同于对知觉或记忆的自信。知觉是对事情的直接定位。知觉的正确性通常被认为是理所当然的。同样，我们每时每刻使用记忆的方式都是直接回忆已经储存在心智中的信息。当知觉和记忆顺利运作、没有阻碍时，我们完全没有意识到它们涉及的推论工作。

但是，我们把直觉视为自身心智的产物，而非从外部环境中或从记忆里提取的信息。我们对自身直觉的自信是对自我能超越给定信息进行推论的心智能力的自信。我们不仅感觉自身的直觉是正确的，还感觉在产生这些直觉方面，我们本人也是正确的。人们认为，直觉推论的结论是个人的想法。当我们想出客观理由来佐证自身直觉时，很容易就假定，必须要将这些客观理由牢记于心，以产生这些直觉。认知能力的这一功能不需要在概念上进行清楚的表达。也许这个功能只是一种元认知性感受，只被激活到让我们关注自身直觉的程度。尽管如此，我们仍然认为，就是这种独具特色的自信体现了直觉的特性。

传统观点认为直觉是特有的直觉官能的产物，当前主流观点认为直觉是系统一类型的推论机制的产物，我们同时摒弃了这两种观点。这就产生了一个疑问：如果谈及直觉自身，直觉既不彼此类似，也不同于其他推论类型，那么我们为何要把它们归为一类，并与其他心理状态区分开呢？我们为何要把这种特殊形式的、使我们与众不同的自信作为直觉，而不把那些没有那么多相似之处的推论当成直觉呢？

为何我们的确有直觉呢？顺便说一句，这是推测的答案。辨别出你自身的某一想法是直觉，便是就这一想法的内容表明个人权威的立场。我们认为，与其说这一立场与你的个人思维相关，倒不如说可能同你与他人交流那一想法的方式相关。直觉是你感知的想法，你可能会为此坚

持自身权威性，不与第三方权威辩论或听取他们的意见。基于直觉做出断言（或提出行动方案）是一种社会性举动，这一举动使他人要么欣然接受，要么表达对你所言及其权威性的不信任。通过表达这种直觉，你在增加赌注：如果这种直觉被接受了，你的权威就提高了；否则，你的权威就降低了。然而，即便你的断言没有被认可，但将其作为你的直觉提出，也可能会帮你反抗权威或他人的论据。直觉会使人更固执，但这有时是一种明智的社会策略。[注]

元认知不仅可以采取评估性感觉这种较为简单的形式，还可以采取元表征性判断这种较为复杂的形式。[注]假如有人问你完成目前这篇论文需要多长时间。你答道："一两天。"但是你没有很强烈的自信。好好想一想，如果你说完成这篇论文"至少要一两天"，你会显得更有自信。这里涉及的比纯粹的元认知性感受要更复杂。你是在元表征两种表征（"我会在一两天内完成这篇论文"与"……至少一两天内"）并比较你对它们的相对自信。这时，元认知也是元表征性的。我们会展示，区分纯粹的元认知性感受与更复杂的元认知性元表征对理解理由极其重要。

我们的直觉源自对无数话题的推论：Bouba和Kiki、摩尔悖论、朋友的情绪或我们可能喜欢哪种类型的电影。而且，我们的元表征能力使我们能形成对自身直觉的直觉。

我们对自身一阶直觉（first-order intuition）的元表征性直觉可能集中在这些直觉的不同方面。例如，它们可能涉及一阶直觉的可靠性，也可能涉及他人对我们乐意与之分享一阶直觉的可接受性。我们的一些元表征性直觉不仅有关我们对自身一阶直觉的自信度，对当前论断至关重要的是，这些元表征性直觉还关于支撑一阶直觉的理由。

你出席了一个派对，而且很高兴在派对上看到了朋友莫莉。然而，她似乎很沮丧。当你找到机会跟她聊两句时，你说道："你今晚似乎很沮丧。"她答道："我并不沮丧。为什么你要这样说？"正如你是凭直觉

知道她很沮丧，现在你也要凭直觉说出支撑你最初直觉的理由。下面便可能是你的两个直觉。

一阶直觉：莫莉很沮丧。

对支撑你一阶直觉的理由的元表征性直觉：“莫莉没有笑且她的声音不自然”，这个事实就是我认为她很沮丧的理由。

你想去看电影，却在纠结该看《超人8》还是《星球大战12》。你凭直觉决定去看《超人8》。结果该电影相当令人失望，因此你反思为何要去看这部电影。你凭直觉得出以下答案。

一阶直觉决定：去看《超人8》。

对支撑你一阶直觉决定的理由的元表征性直觉，“比起《星球大战11》，我更喜欢看《超人7》”，这个事实是你决定去看《超人8》而非《星球大战12》的理由。

当自身直觉遭到他人质疑或被之后的经历动摇时，我们通常会比较关心自己的理由。

我们不仅可以凭直觉知道支撑自身直觉的理由，还可以凭直觉知道他人支撑其自身直觉的理由。你在一个没有窗户的会议室里待了数小时，现在你与同事小林（Lin）一起走出了这栋楼。天很蓝，和你进来的时候一样，然而小林却说“下过雨”。他这么说的理由是什么呢？是的，气温的确比想象中冷一点，但这足以佐证小林的直觉吗？你环顾四周，发现了一些水坑。你凭直觉知道“有水坑”为小林的断言提供了理由。

小林的直觉：下过雨。

对小林支撑其直觉的理由的元表征性直觉：“有水坑”，该事实是小林猜想下过雨的理由。

此外，如果你认为你归于小林的理由很充分，就不必只是因为其权威而接受其断言，现在你也提供了接受其断言的理由。

理性模块

让我们能凭直觉推断出理由的这种机制是什么？让我们最初便能像这样识别理由的实证规律又是什么？虽然理由的归因很难讨论，但信念及欲望的归因一直是哲学、心理学辩论的中心议题。正如我们所见，普遍的观点认为，使我们把信念及意图归于人的这种规律便是对人类理性的假定，我们对这一观点持保留意见。对理性的这种假定难道不能帮我们识别人们的理由吗？人们应该把使其信念及意图变得理性的理由归结于他人吗？

那么我们自身的理由又如何呢？我们如何得知自身的理由，或以为我们知道？识别他人及我们自身的理由是基于同样的机制，还是基于两种运行方式截然不同的机制呢？

正如我们在第七章中所知，有人认为自省的天赋或力量允许我们直接解读自身心智，对此我们有充分的理由可以不予接受。此外，我们不是在得出自身直觉结论时想出理由并对这些理由进行自省的——前提是有自省这回事。相反，我们通常是在得出结论后才构想支撑该结论的理由。正如我们必须推断出归于他人的理由一样，我们也必须推断出归于自身的理由。无疑，我们对自身通常会有许多更充足的证据。我们在某种程度上能直接获知感觉与感受。我们能用内部语言与自我交流。尽管可用证据存在差异，但在至关重要的方面，我们推断出自身理由的方式应与推断出他人理由的方式类似。

这是否意味着，当我们把理由归于自身或他人时，必须假定自身（或他人）是理性的且往往会得出有充分理由支撑的结论，并以此为一个前提呢？不，从模块主义的观点来看，规律（如人类倾向于理性地思考与行动这一规律）能被模块程序所利用，而根本不以被表征为前提。换句话说，我们可以利用人类理性的特性，把理由归于他人或自身，而不必对人类理性本身持有任何普遍看法。

在搜寻理由时让理性发挥主要作用，这一想法尽管有吸引力，但仍有许多问题存在（不论是以标准视角还是以模块主义视角来看）。突出理由与理性间的紧密联系很有吸引力。如果我们不理性，就不会理解理由，更不用说将其放在心上，这一点似乎符合基本常识。如果理由在某种程度上不是理性的，我们甚至不会将其视为理由。

更成问题的是，有人认为人类倾向于理性地思考的普遍趋势（连同每个案例的具体信息），应该足以指导我们识别理由。无数截然不同的理由被用来支撑任何理性的信念或意图，而且所有这些理由都与可用的有限证据相辅相成，并共同做出理性的辩护。

或许小林做出“下过雨”的断言并不完全是因为看到了这些水坑，还可能是因为他有识别并解读空气温度及湿度变化的卓越能力；当你们都在会议室时，他可能注意到有人拿着一把湿伞进了门；或者他可能有截然不同的理由，而你们对此毫无头绪。这表明，仅仅寻找那人可得到的客观理由，还不能与发现那人真实理由的有效启发法相提并论。

我们在第六章中谈论“读心”时，遇到了类似的问题。理性不能单独为信念及意图归因提供充分依据。我们认为这种心理状态归因利用了心智的模块组织。这一模块组织利用的不是作为普遍特点的理性，而是人类共有的各种认知能力的具体特点（这些具体特点共同促使人类变得理性）。我们现在认为，理由归因也是如此。要把理由归于他人或自身，人类更多的是依靠某些具体能力的效用，而非自己全部的理性。

回到派对上莫莉的那个例子。她质疑你，让你解释为何认为她很沮丧。为了给出回答，你必须做一个逆向推论，从你最初对她情绪的直觉推出会佐证你直觉的理由。你不确定是什么触发了你的这一直觉，但是你可以做出一些貌似可信的猜测。

你不仅是一个理性的人，你还具备人类特有的情绪辨别能力。你能用面部表情、语调、肢体动作等作为情绪证据来“解读”他人情绪，这是你的专长。虽然你记不住让你有此直觉的认知过程（直觉推论过程本来就很晦涩），但你记住了遇见莫莉当晚注意到的一些事情，那些便是情绪的标准指示器：她没有笑、她的语调不自然等。所以，你从记得的（或者说从你当前正注意到的）一些事情中提取碎片化的证据，使其能完美地佐证你认为莫莉很沮丧的直觉，接着你便推断出这些就是支撑你直觉的理由。

通常，你对某人情绪的直觉是由各种因素共同触发的，而且有许多因素你甚至都没有意识到。在所有因素并存的情况下，每个因素都会对你的直觉有一定的影响。当你选取其中一些证据作为你直觉的理由支撑时，通常会夸大其作为证据的分量，但此举可能便是明确的叙述理应具备的条件。

正如我们从第三章中所知，你的记忆不仅是对过往的回忆，还具有建设性的意义，且通常会“记住”那些有助于更好地了解所发生之事的特征，即便你当时并没有注意到这些特征。例如，你可能并没有注意到莫莉的语调不自然，直到她问你为何认为她很沮丧。尽管如此，因为它与你认为莫莉很沮丧的直觉非常相称，所以这一特征在你思考最初是什么促成了你的直觉时，就注入你的记忆中。你现在援引的这些理由的说服力本身就是推断出来的。它是从你对自身直觉的自信中推断的：如果你感觉自身直觉是正确的，那么支撑这一直觉的理由就一定是充分的。你寻找貌似可信的充分理由，并假定它们就是刺激过你的理由。

我们凭直觉推断出自身支撑某一特定直觉的理由，并不是基于对自

身理性的一般推测，而是基于对产生这一直觉的某种具体能力更狭隘的自信。当我们凭直觉推断出某个朋友的情绪时，我们之所以感觉自己是正确的，不是因为理性，而是因为我们有能力去判断人们的情绪，特别是很了解的人的情绪。

当你必须推断出导致另一个人得出给定结论的理由时，你的任务还是做一个逆向推论，从这一结论推出能解释并至少在某种程度上佐证它的理由。注当你与小林一起走出大楼时，他说“下过雨”，你环顾四周，看到了一些水坑，你假定这些水坑的存在为小林的断言提供了充分的理由。为何你能凭直觉知道水坑是推断出下过雨的理由？因为，就像小林一样，你有能力去识别这个有关天气的证据。当小林的陈述促使你去关注这些能说明问题的水坑时，你自己便能凭直觉知道肯定下过雨。既然你相信自身的推论，即从水坑推断出下过雨，所以你便做出更高层级的推论，即认为这种水坑的存在对他人或你自己，都是推断刚下过雨的充分理由，接着你便进一步推断这很可能就是小林的理由。

我们不能理解（共享）他人直觉时又该如何呢？假设当小林说下过雨时，你看到了水坑，但不能凭直觉知道必然下过雨，而只是推测可能下过雨。实际上，再仔细观察，你会注意到这些水坑（并没有太多）都在你面前同一块有限的区域内。现在你更倾向于推断出，只有你面前的这块地面碰巧被水淋过了，它并不是被雨淋的，因此没有下过雨。尽管如此，这些水坑的确向你展现了下过雨的可能性：这是推断出可能下过雨的表面证据，却不是充分的理由。因此，你的直觉仍然会是，小林断言“下过雨”的理由很可能是他注意到了这些水坑，但你现在会判定这是一个不充分的理由。另外，如果你没有想到任何下过雨的证据，那么可能会对小林的理由没有直觉。你可能只是对他的断言表示困惑并一直困惑。

把理由归于他人或自己，不过是一个相当浅显的事情。人们在环境或记忆中搜寻一些真实的或貌似可信的信息（莫莉不自然的声音、水

坑），这些信息可以用来解释并佐证直觉。如果找到了这样的信息，那么，这便被视作支撑某人自身直觉的真实理由、支撑他人直觉的极有可能的理由。

与常识性观念相反，我们并不是从以某种方法获得的理由中推导出直觉结论的，而是通过进一步的直觉逆向推论过程，从这些直觉本身推导出支撑我们直觉的理由。我们从自己凭直觉得出的结论中推断自身必然持有的理由。我们通常构造理由来作为事后的合理化。

我们以同样的方式把理由归于他人：基于对他人能力的信任程度来信任其直觉，并通过同样的逆向推论过程来推断其理由。当我们不信任他们的能力时，类似的逆向推论过程会勉强形成一些理由，我们发现这些理由太无力或漏洞太大，根本不能佐证他们的直觉结论，但他们自己可能会觉得理由充分。当我们认定别人犯了错时，通常乐于把一些明显很糟糕的理由归于他们。

我们推断出自身的理由，使之支撑自身的直觉结论。我们基于对他人结论的认同程度来评估其理由的力度。这是否意味着，理由的搜寻是一个纯粹表面的事情、一种掩饰我们赤裸裸偏见的方式，只为了让自己看起来理由充分，而他人的理由看起来充分与否则取决于我们是否同意其观点呢？我们从这一过程中得不到认知利益吗？若得不到，这一过程就不会有太大意义。如果不论发生什么，每个人只是一味坚持他们起初的直觉，理由就会变得完全无关紧要了。

我们证实过，理由是为了社会消费。人们想出理由来为自己做解释或辩护。如果这样做，他们就承担了其所辩护的想法和行动的责任，隐性地遵守一些他们认为合理的规范并希望他人也遵守。人们在给出理由时，便冒着其理由会被质疑的风险。他们也有权利去质疑他人的理由。在很大程度上，某人的声誉便是关于这个人的理由的会话在时间和社会空间中延续的持续影响。我们给出自己的理由，就是为了参与到谈论自己的会话中并捍卫自身声誉，通过评估并讨论他人理由的方式来影响其

声誉。[注]

所以，我们援引理由不是为了无意义的自我吹捧，而是为了让推断自身理由的方式偏向于对自身有利的一面。我们想要用自己的理由来为自己在他人眼里的形象辩护，因为这些理由会供他人判断，会被重新考虑，然后被修改以便更好地被接受。而且，这有时意味着要修改我们的理由所支撑的结论：改变观点或行动方案以更好地为自己辩护。最终，理由和结论可能需要再调整。

我们假定存在一个专门的元表征模块，其任务是推断我们自身及他人的理由。它的任务不是在心理学上准确解释刺激人们的理由。实际上，内隐心理学（即假定人们的信念和行动受到理由的刺激）经实验证明是错误的。给出理由来为自身辩护并对他人给出的理由做出反应，是塑造声誉和协同期望的一种方式。

这是否意味着我们给出的理由及希望他人给出的理由只是调整一些局部共识、一些有关理性的文化构造概念？是否意味着我们不应该预期人们的理由从客观的角度来讲是理性的？不，根本不是这个意思。人们所给的理由不仅表明他们是社会族群中遵守规范的成员，而且发挥着更加重要的作用。当人们被视为可靠的信息来源及有用的合作伙伴时，他们就获得了他们在意的良好声誉。随着时间的流逝，如果没有基本的客观理性来使他们做出认知上合理的推论并使之采取有效行动，他们就没有办法维持这样的声誉。为了自己的声誉着想，人们援引的个人理由应该被他人认可为客观理由的表征，获得这一认可的最好方式至少是援引客观上合理或至少不太荒谬的理由。

文化社群可能会倾向于几种明确的理由，如对特定权威的依赖。对于不同类型的人，如女性和男性、年轻人和老年人、社会上的强者和弱者，文化社群不会选择性地认可他们援引理由的能力，也会谅解一些不理性的理由，如有预兆的梦。社群不能做的便是构造一连串自己的理

由。无论在哪里，人们对理由的直觉都根植于认知能力，这种认知能力在很大程度上是他们作为人类一员所共有的，且有助于提高人类的认知效率，从该术语的基本含义来看，就是理性。没有这种认知基础，我们不相信任何理性规范都可以在社群中诞生并被保留。

我们的理由往往是理性的，因为一开始我们的直觉往往是理性的。人类所做的，且其他动物大概不会做的，便是向其无意识推论中加入对支撑这些推论的理由的更高层级的表征。从与生物学相关的认知效率来看，这些理由并不能使人类的推论理性化。

这些理由反而有助于对我们的推论进行理性的表征——若取该术语另一种与社会相关联的意义，在这层意义中，准确来说，有理性就意味着以能清楚表达并评估的个人理由为基础。[注]个人理由指导下的信念及意图的公共表征是人类社会互动的基本层面。我们认为这些表征产生于专门的元表征性模块。我们所有的理由——不论是直接的还是间接的——都是这一模块的产物。

那么，人类理性是一个模块吗？我们发现这一模块了吗？下一步应该是确定理性模块在大脑中的位置吗？不，不是的。传统的理性观念并不是关于心理学机制的，而是关于人类心智作为一个整体本质的、先验的特点。总之，我们还没有找到任何模块；我们希望，我们只是在用合理的论证推测出，识别理由很可能是某个特定模块的任务。如果我们的推测正确且存在这样一个模块，该模块的确会在某一神经结构中被发现，但那一结构不需要在大脑中占据（且单独占据）任何位点。总之，这样一个用来推断理由的模块不会与传统意义上的理性相一致。

笛卡儿等许多哲学家都在为理性高唱赞歌（尽管其他思想家并没有那样予以热情赞美）。我们正谈论的这个模块如果存在，人类是不会像吹嘘理性那样去吹嘘该模块的。尽管如此，在人类大脑中能找到的最接近传统理性的东西很可能就是这个模块了。因此我们会把它称为理性模

块。

理性模块能推理吗

至少在某种程度上，理性模块旨在产生正当理由，且非常偏向于对我们有利的一面。如果产生于这一模块的理由是通过逆向推论从直觉推论中推断出来，那么要如何改善我们的直觉推论呢？逆向推论在推理中有什么用呢？我们对他人理由的评价是如何凌驾于自私自利的偏见之上的？我们如何被他人的理由说服，进而改变自己的观点？

部分答案在于我们的一阶直觉（莫莉的情绪、下雨以及使我们有此直觉的各种事情）是由许多模块产生的——尽管对支撑一阶直觉的理由的元表征性直觉产生于一个只作用于理由的元表征性模块。在其能力范围内，一阶模块利用其领域的规律做出了有关客体的各种推论。所涉及的元表征性模块在其能力范围内做出推论，即做出关于理由和结论间关系的推论。若这样做，该模块便注意到了这一关系的相关特性。我们很容易出现一些直觉，且感觉其非常有力，但是发现与其本身同样有力的理由并没有那么容易找到。对支撑某一直觉的理由缺乏信心会渐渐破坏最初对那一直觉的满满信心。

例如，我们天生倾向于去承担难得机遇所伴随的重要风险。这样一种倾向在过去通常可能是有利的，而在现代社会，却容易被骗子利用。

例如，杰布（Jeb）产生一个瞬间直觉，如果他积极回应刚收到的这条来自富有银行家遗孀的信息，即让杰布帮她转移几百万美元的资产到他的国家，那么杰布便会变得极其富有，但紧接着他可能很难找到可信的理由——能与亲友分享的理由——来支撑这一直觉。朋友妮娜（Nina）警告他这是一个骗局，虽然他最初会对此不予理会，但很快便会找到支撑它的充分理由。他对理由持怀疑态度的直觉不同于且好于他

对大笔财富热情洋溢的直觉。

他人通常会表达不同于我们的直觉。当我们试图去解释这些自己并不具有的直觉时，会凭直觉把自己发现的充分甚至强有力的理由归于他们，进而促使我们修改自己最初的直觉。尼科斯（Nikos）和索菲娅（Sofia）的一个朋友让他们解决下面这个问题（实际上，推理心理学最近对该问题做过很多研究）[注]。球拍和球一共值1.1欧元，球拍比球贵1欧元，那球多少钱呢？尼科斯的直觉是球拍值0.1欧元。让他感到惊讶的是，索菲娅却答道：“球值5分钱。”看到他很困惑，索菲娅继续说道：“如果球值5分钱，那么球拍必须是1.05欧元……”在她说完之前，尼科斯明白了：它们的差是1欧元，刚好符合要求！他凭直觉知道这一定是索菲娅支撑其答案的理由，而且是一个充分的理由。因此，他抛弃了自己最初认为球值0.1欧元的直觉，接受了索菲娅的答案。

我们一阶直觉的力量与相对应的元表征性直觉的力量不需要相称。理由可能会加强或减弱我们的一阶直觉，且有时还会促使我们对此进行修改。这样，理由便不必成为认同我们一阶直觉的纯粹标志。

不过，不要把杰布或尼科斯的例子当作典型。杰布最初的直觉并不符合常识。很明显，如果他去做了，就必须要证明此举的正当性，但他很难找到貌似可信的正当理由。之后他得到了妮娜的帮助，找到了让他重新考虑的理由。同样，对于球拍和球这个问题，如果索菲娅没有给出不同的答案且进行解释，那么尼科斯可能不会修改他的答案。通常，我们的直觉既不会像杰布那样愚蠢至极，也不会像尼科斯那样错得离谱。我们大多数人的一阶直觉至少是可信的，且逆向推论通常会用貌似有理的二阶理由来佐证它们。

除非我们的声誉危如累卵，否则我们不可能根据对理由的元表征性直觉来认真检查自身的一阶直觉。即使我们认真检查了自身的一阶直觉，很容易就能想到的理由很可能会证实甚至加强最初的直觉。即使我

们的一阶直觉和对理由的二阶直觉不一致，我们也不会不假思索地更相信后者。像杰布这样的骗局受害者，可能会有一些更高层级的疑虑，但仍会落入圈套，因为他的一阶直觉太有力了。所以，一阶和二阶直觉可能不一致，但这一事实还不足以使我们变得聪明睿智。

我们带着这个双重谜团开始写这本书，本书的第二部分是：人类为什么不更擅长推理呢？为什么通过推理也不能达到近乎普遍的一致呢？现在看来，我们可能过度解释了为何不同人的理由不能汇集达成同一个结论。在本章结尾，我们提出了一个相反的问题：如果理性模块适用于回顾性运用理由来做辩护，那么又如何能被用来进行预期性推理呢？人类如何能在任何情况下都有推理能力且有时推理得非常好呢？

---

1. Raz 2000, 2010;也可参见：Wallace et al. 2004。
2. Carroll 1895。更多丰富的现代论述，与心理学相关的知识，可参见：Kornblith 2012；Railton 2004。
3. 可参见：Köhler 1930; Ramachandram & Hubbard 2001。欲了解最近对该问题的相关研究，可参见：Monaghan et al. 2014。
4. Green & Williams 2007.
5. 我们在第三章中论证的问题，也可参见：Thompson 2014。
6. 欲找到优质的例子，可阅读罗伯特·泽拉塔斯基（Robert Zaretsky）和约翰·斯科特（John Scott）合著的《哲学家之争》（*The Philosopher's Quarrel*）(2009)，其中，他们描述了让-雅克·卢梭(Jean-Jacques Rousseau)是如何凭借自身直觉固执地反对大卫·休谟明智的论据的，支撑其直觉的是他自己的哲学论据。
7. 详细论证可参见：Proust 2013。
8. Mercier 2012.
9. 一般性声誉问题可参见：Origgi 2015。道德性声誉问题可参见：Baumard, André & Sperber 2013; Sperber & Baumard 2012。
10. 这一理论与尤尔根·哈贝马斯的交往行为理论（Jürgen Habermasd，1987）的符合程度值得探究。
11. Frederick 2005.

第九章
推理：直觉和反思

理由能佐证某个已形成的观点或已做出的决定，这便是理由的回顾性运用。但是如果你心中有不知道该如何解答的疑问，或者不知道如何决定，又该怎么办呢？这时，理由的预期性运用——合理推理——不是能帮上忙吗？原则上是的。如果你对该问题的直觉不够明确或直觉不够有力，以至于你左右摇摆，那么用一种不偏不倚的方式来思考支撑正确答案或最好决定的理由，似乎再好不过了。这是理性模块无论如何都会帮你做到的吗？

直觉论据，反思结论

我们已经强调过，理性模块提供的直觉不是关于可以成为支撑某一非特定结论的理由的事实，它们的形式不是“P是一个理由”（例如，“艾米发烧了是一个理由”）。这些直觉与一些事实及其支撑的结论相关，其形式是“P是支撑Q的理由”（例如，“艾米发烧了是打电话叫医生的理由”）。

在辩护和推理中，理性模块的产物都是高阶结论，即存在支撑某低阶结论的理由。就辩护而言，低阶结论已经得出且可能已被遵照，即艾米的父母已打电话叫了医生。现在补充的是高阶结论，即艾米发烧这个事实已佐证了这一决定。即使艾米的父母在实际给她量体温之前打电话叫医生只是因为她头晕，这一回顾性辩护也会行得通。

在推理中，理性模块产生的不是一个而是两个新的结论，第二个结论隐藏在第一个结论中。第一个结论本身就是高阶论据，即元表征性直觉——“诸如此类的理由支撑某一特定结论”：艾米发烧是现在就打电话叫医生的充分理由。第二个新结论——我们打电话叫医生吧！——是隐藏在整个论点中并由其支撑的。

如果你抓了一条刚整个吞下另一条鱼的鱼，那么你就一次抓住了两条鱼。同样，如果你推断出的某一新结论受到充分理由的支撑，那么你便一次推断出两个新结论，即整个论点及其支撑的结论。当隐藏的结论本身就很中肯时——因为它在推理中很常见，你会将其取出并维护它，而后存储在记忆中，或在将来的推理中用作前提。

以下有两个实例。

你十分确定把所需的一本书落在书房了，但无论你多努力地寻找，都没能找到。你尽力去想还可能放在哪些地方了，然后你想到，也可能会在起居室里。这时，你以论点的形式得出了一个结论。

> 论点：既然我十分确定书要么在书房，要么在起居室，且我在书房没找到，那么我应该去起居室再找找。

这便是人们通常会接受的论点，而没有考虑支撑它的更高层次的理由。简单而言，直觉可以定义为人们在未觉察甚至未意识到支撑性理由时便接受的结论，鉴于此，你的论证从整体上看无疑是一个直觉结论，一种直觉。不过，这种直觉结论是关于理由的，是关于这些理由所给予第二个结论的支撑的，这第二个结论便隐藏在下列论点中。

> 隐藏的结论：我应该在起居室里找找那本书。

“你应该在起居室里找找那本书”这个结论也是一种直觉吗？不，因为它并不符合直觉的定义。毕竟，它是有理由支撑的，那些你铭记于心的理由。实际上，这一结论隐藏在对这些理由的表征中。

如果隐藏在论证中的结论不是直觉结论，那么它是哪种结论呢？它是经过推理的结论，或者说是反思结论，这是推理文献中通用的术语，我们自己也经常使用，注接受该结论是因为对它进行了高阶思考（或“反思”）。正如这个例子阐明的那样，得出反思结论所用的反思可能是最少的。尽管如此，“不用进行任何推理就凭直觉认为应该去起居室找这本书”和“有支撑该结论的直觉理由”之间依然存在差异。注

下面我们会看到第二个关于反思结论的例子，这次只需要多一点点反思。你和一个朋友坐在咖啡厅里，他向你发起挑战：“我给你两个赌注，接受其中一个，如果你赢了，我就埋单。”他给的两个赌注如下。

赌注一：如果接下来进入咖啡厅的五个人中至少有三个人是男性，你就赢了。

赌注二：如果接下来进入咖啡厅的六个人中至少有三个人是男性，你就赢了。

不用考虑高阶理由，你想到了一个能帮你做出选择的直觉论据。

论据：赢得赌注二比赢得赌注一的概率要大，选择赌注二更好些。

隐藏的结论：赌注二是更好的选择。

你对赢得这两个赌注相对概率的直觉，支撑你得出“赌注二更有利”这个结论。因此，这一直觉提供了支撑反思结论的理由。

这两个例子展现的是，反思结论不必成为可以与直觉推论机制形成对照的反思推论机制的产物。反思结论可能是直觉推论过程的间接产物。这一过程的直接产物是对支撑某结论C的理由R_1、R_2……R_n的直觉结论——论据，间接产物是反思结论C。可参见表9.1。

表9.1 理性模块的产物

	结论类型	结论形式
直接产物	直觉论据	支撑结论 C 的理由 R_1、R_2……R_n
间接产物	反思结论	结论 C

我们想要更强有力地强调：在人类思维中，全部的反思结论都是对理由进行直觉推论的机制的间接产物。

在基本的和最普遍的案例（例如，错放的书或刚才所述的对赌注的选择）中，反思结论隐藏在直觉论据中。在更复杂的案例中，可能存在两个隐藏层次而不是一个，直觉论据支撑反思论据，反思论据又支撑最终的反思结论。

回到赌注这个例子。我们把你描述成在直觉上认为赢得赌注二比赢得赌注一的概率更大。这一直觉为你选择赌注二提供了强有力的论据。不过，你可能会进一步反思，然后仅基于更高层次的直觉论据便接受这个支撑赌注二的论据。

> 直觉论据：如果接下来五个进入咖啡厅的人中只有两个人是男性，你会输掉赌注一，但仍会赢得赌注二。只需要第六个进入咖啡厅的人是男性就行了。因此，赢得赌注二比赢得赌注一的概率要大，这便使赌注二成为更好的选择。
>
> 反思论据：赢得赌注二比赢得赌注一的概率要大，这便使赌注二成为更好的选择。

反思结论：赌注二是更好的选择。

在这个例子中，对大多数人而言，更高层次的论据（其中反思论据和最终结论都隐藏了）很可能是直觉性的。有些人不认为这一更普遍的论据是直觉性的，有些人不愿不经过反思就接受论据，对他们来说，更高层次的论据本身可能隐藏在会支撑该论据的、更普遍的概率论据中。可能存在两个以上的隐藏层次。尽管如此，最终任何反思结论都是直觉推论过程的间接产物。

认为高阶显性论据使结论更容易被理解和接受是一种错误。举个极端的例子，对大多数人而言，“1+1=2”是很明显、无可辩驳的事实。众所周知，阿尔弗雷德·诺思·怀特海（Alfred North Whitehead）和伯特兰·罗素（Bertrand Russell）在他们里程碑式的专著《数学原理》（*Principia Mathematica*）[注]中，通过运用一连串的复杂论据，用了好几百页纸，只为获得这一结论，只有极少数人能看懂这些论据。总之，这些论据并不是为了强化这一坚如磐石的普遍直觉，即1+1=2，而是表明该直觉是可以证明的，并且这样做会为数学提供逻辑根据。

人们从生活中得出的许多结论中，哪些是反思而非直觉的呢？这些反思结论本身又有哪些隐藏在反思论据而非直觉论据中呢？更高层次的反思与高阶反思有多普遍呢？这些都是还没有被充分研究过的实证性问题。

值得一提的是，逻辑学家、哲学家和推理心理学家最可能带着专家的自信表达自身的观点，指出反思在人类推论中的重要性。这些专家一般并不能作为全人类的典型；他们通常诉诸更高层次的论据，并将之作为其职业的一部分。因此，他们可能容易将自己职业性的心理扭曲，误认为是人类的基本特点。我们猜测，人类大多数推理——即便是那种本应使我们十分优越的精彩推理——也很少涉及超过一两个层次的论据。

我们在第八章中已经表明，理性模块作为直接产物，产生对理由的直觉，作为间接产物，得出由这些理由支撑的反思结论。如果这是正确的，那么便没有必要——实际上没有机会——运行心理学机制（“系统二”或“系统一”），该心理学机制的任务是直接产生反思结论。推理在多大程度上是反思性的，取决于所涉论据本身隐藏在高阶论据中的程度。不过，更多的反思性并不意味着更少的直觉性。从直接隐藏在直觉论据中的初级反思结论到一些最引人注目的科学成就的多层次反思特点，推理总是直觉推论机制的产物。

当把推理描述为对理由的直觉运用时，我们采取了一种纯粹心理学的方式，这一方式与较传统的逻辑推理方式大相径庭。难道逻辑不在推理中发挥主要作用吗？难道我们没有错过某些极其重要的事情吗？

为了更好地理解这个问题，要把推理心理学与数学心理学进行比较。数学科学本身讲述的就是客观的数学事实。数学心理学讲述了人们学习这些事实并利用它们来计算的方式，例如，计算应付利息或花园的表面积。我们不能用纯粹心理学的方式来看待数学心理学，从数学事实只能窥豹一斑。同样，据称推理便是对客观逻辑事实的运用。对理由的直觉无关理由的逻辑性质，不属于推理，正如对幸运数字的信念也不属于数学思维一样。如果这一论点正确，那么推理心理学的核心应该是逻辑。

数学能力的习得及运用依赖表现数字、运算等的数学符号。甚至在发明书写之前，许多口头语中使用的词语都被用作数学符号，且能够轻微地促进数学能力的发展。同样，逻辑依赖符号来表征命题、逻辑关系等。直到最近，一套完整的专用逻辑符号才得到运用，但就像在数学中一样，日常语言的许多词语及表达都能作为逻辑符号。所以有人认为，在更形式化的逻辑得到发展之前，语言便促成了推理。

所以，在传统推理方式中，语言对推理必不可少。我们认同该观点，却出于完全不同的理由。

推理依赖语言

我们主张，理由是为了社会消费。为了社会共享，理由必须用言语来表达，的确，理由都是以语言形式出现在思维场景或公共场合中。理由可以用来为自身做辩护、评估他人或去说服想法不一致的人。所有这些都涉及语言沟通。(注)即便当你独自思考理由时，通常情况是，你仿佛在心里回答他人所说或可能会说的话，仿佛在让自己准备好去质疑他人的想法或举动。即便当你思考理由来回答自身疑问时，也仿佛让自己置身于与自我的对话中。为此，你求助于内部语言。

只有语言能很好地表征理由。要理解理由便是在心理上表征至少两种表征之间的联系，即理由本身和其支撑的结论；换句话说，它是一个元表征性任务。语言和元表征联系紧密（即便没有语言，元表征尤其是"读心"过程也行得通）。(注)语言是清楚表达复杂元表征并就此进行沟通的独特而高效的工具。尤其是语言，特别适合于元表征理由和结论间的联系。

语言表达能嵌套在语言表达内，特别是句子能嵌套在句子里："它很好"是一个句子，"雅思米娜（Yasmina）说它很好"也是一个句子。好几个句子能嵌套在一个更复杂的句子里来清楚地表达理由。例如，

莫莉没有笑。

莫莉很沮丧。

这是可以表征事态的句子。这些句子能结合在一起，就像"莫莉没有笑"这个事实是推出"她很沮丧"的理由。

这个复杂的句子元表征了理由和其支撑的结论间的这种联系。

为展现各种元表征性的联系，特别是理由-结论间的联系，语言提供了各种各样的语言策略。正如“论据”“理由”“异议”“结论”这些名词能描述状态及其联系，“辩论”“反对”“推断”这些动词也能。许多所谓对话标志，诸如连接词“所以”“因此”“尽管”“但是”“即使”“然而”，都有辩论的功能：它们关注理由-结论关系，而且就其特点给出一些指示。(注)

例如，对问题“派对怎么样？”，请比较以下两个回答。

回答一：派对还不错。巴勃罗（Pablo）带了他的尤克里里琴。

回答二：除了巴勃罗带了他的尤克里里琴之外，派对还不错。

虽然两个回答都陈述了同样的事实，却被置于不同的视角之下。在回答一中，“巴勃罗带了尤克里里琴”这个信息会被理解为“派对还不错”这个陈述的细化和证明；音乐促成了派对的成功。在回答二中，相反，“除了……之外”暗示，你从“派对还不错”这个断言中可能会推断出的某些结论，如“所有事情进展顺利”，实际上并不符合这种情况。是的，派对还不错，但是“尽管有”而不是“因为有”巴勃罗的音乐。

辩论策略，如“除了……之外”这类词，在论证中发挥着启发式作用。它们能促进推论，并显示哪一种暗示应该或不应该从上下文中获得。不管怎样，根据传统观念，在语言推理中，以语言形式表达的逻辑符号发挥着主要作用，特别是逻辑术语，如“或者”“如果……那么”“只有……才”“并且”和“不”（及其他的逻辑策略，如数量词和情态动词）。这些逻辑策略有助于构造有效的演绎论据，是促成推理（或至少是所谓演绎推理）的语言工具。真的如此吗？逻辑策略是推理的主角，而其他语言策略只是配角？接下来，我们要讲述一个不同的故事。

奇遇——被偷的钻石和丢失的前提

与公认的观点相反，合理三段论（及其他合理的正式演绎）在日常推理中并不能提供令人信服的论据。我们用夏洛克·福尔摩斯（Sherlock Holmes）的一段新奇遇来阐明为何不能。

夏洛克·福尔摩斯收到一封电报。他读完后便给朋友华生（Watson）医生回信：

福尔摩斯：男管家没有偷钻石。所以，是园丁偷了。

华生出于对福尔摩斯演绎推理能力的信任，可能会被说服。

不过，从标准的逻辑学观点来看，福尔摩斯没能展现正当的演绎论据。请注意，他使用了辩论策略词“所以”，而没有使用逻辑术语。他的陈述充其量只是一个省略三段论，即真实逻辑论证的缩减版本。如果是这样，福尔摩斯本可以更直截了当些。

福尔摩斯：男管家或者园丁偷了钻石（第一个前提）。不是男管家偷的（第二个前提）。所以是园丁偷的（结论）。

这个使用了逻辑术语“或者”和“不是”的逻辑版本与析取三段论（disjunctive syllogism）相对应。

前提：1. P或者Q。

2. 不是P。

结论：是Q。

这种三段论可以认为有正当根据。当有正当根据的三段论的前提正确时，该三段论有效且合理：合理三段论的结论必然是正确的。也许华

生认可福尔摩斯以缩减形式表现的完整三段论并心悦诚服，不仅出于他信任福尔摩斯对逻辑的敏锐，而且出于该论证的真实逻辑。

然而，你所知的合理三段论本身给你提供了支撑其结论的充足论据吗？这样想便犯了普遍性错误。

假设福尔摩斯刚收到的这封电报是警察局长发的，写的是：钻石是园丁偷的。在那之前，福尔摩斯都不知道钻石可能是谁偷的，恰好就认可这是事实，但他没有直接告诉华生，而是决定用三段论形式来表述这个信息。确切地说，他可能做如下表述。

福尔摩斯：教皇或者园丁偷了钻石。不是教皇偷的。所以是园丁偷的。

福尔摩斯的伪论证明显是在绕圈子，华生应该能轻易辨识：既然教皇从不会被认定为罪犯，福尔摩斯论证的第一个前提（“教皇或者园丁偷了钻石”）唯一貌似可信的根据便是他已经知道是园丁偷了钻石，而且他本可以直接这样说。三段论显然是合理的，但华生不会认为这是真实的论据，而会认为这是福尔摩斯用来表述自己已知事情的一种古怪的方式。这种情况下，华生接受“是园丁偷了钻石”的理由与福尔摩斯伪论证的逻辑合理性没有关系，却与他对福尔摩斯的信任大有关系。

在同样的条件下，“狡猾的”福尔摩斯本可以说出合乎逻辑、十分真实但具有误导性的话。他本可以这样说：

福尔摩斯：男管家或者园丁偷了钻石。不是男管家偷的。所以是园丁偷的。

把教皇换成男管家并没有改变三段论的一丁点儿逻辑，也没有改变

其合理性，但很可能诱导华生把它当作真实的论据、接受其结论的理由。为什么呢？因为，不像教皇，男管家是可以被怀疑的，而且福尔摩斯可以从“男管家没有偷钻石”这个信息演绎出“是园丁偷了钻石”。华生可能会认为，福尔摩斯不仅拿出了合情合理的三段论，而且拿出了真实的论据。

更为普遍的是，对于任何你碰巧知道的事实，你总能构造出把这一事实作为结论的合理三段论，并表明你能知道这个事实多亏了自身的演绎能力。然而，这种三段论不能给其结论任何真实的支撑。要使三段论作为真实论证以支撑其结论，必须有合理的非循环论证的理由来认可其前提。这不是逻辑上的需求。从逻辑学的观点来看，循环论证不是问题，可非循环论证却是优秀推理的要求。因此，当完整三段论独自作为论据时，这个论据本身便是不完整的；它包含隐性的前提。它暗指有合情理、非循环论证的理由来认可显性的前提，最终接受其结论。

普遍的观点认为，常用在沟通中的大多数论证实际上是缩减的三段论（有时如此，但比通常假设的要少得多）。我们刚才展示的是，只由三段论构成，即便是完整显性的三段论构成的论据，其本身便是缩减的论据。它们隐性地表达，其前提不仅是真实的，而且还提供了接受结论的真正理由。

在推理中，逻辑是启发式工具

对探究式或论证式推理而言，三段论并不是较好的论证。它们完全是不同种类的事物。[注]三段论（通常还有演绎推理）是抽象的形式化的或半形式化的结构，在前提和结论的逻辑结果间建立了清楚明确的联系。它们可以有各种用途，而自身却什么都不做。但是，在推理中用到的论证，并不是根据其变幻莫测的结构来定义的，而是通过其所做的事情，即给推理者提供得出某一结论的理由。

仅承认三段论是有效的并不是认可其结论的理由，这一点司空见惯。而且，我们也已展示了为何承认三段论的有效性不是认可其结论的充分理由。只有当你有单独的理由来认可其前提时，合理三段论才会给你认可其结论的理由。在这种情况下，三段论所做的便是让你看清支撑前提的理由是如何同样支撑结论的。

如果你认为某个给定的三段论是合理的，且的确有充分、非循环论证的理由来认可其前提，将会怎样？在这种情况下，难道你不应该认可其结论吗？答案还是否定的。你现在拥有的是认可其结论或改变自身想法的理由，至少是支撑其中某个前提的理由。这远非纯粹的理论可能性或稀缺性。实际上，在推理中普遍运用三段论便是通过向人们展示其信念中蕴含着强有力且独立的理由来反驳的结果，从而促使他们修改自身信念。

归谬法（reductio ad absurdum）作为一种论证形式，表现为通过显示某种断言会导致荒谬性或至少显然的谬误来驳斥它。[注]假设华生在看到了福尔摩斯对“园丁是小偷”的表象论证后，回电报告知他：“亲爱的福尔摩斯，你的三段论像往常一样没有瑕疵，但你的结论不可能正确：园丁心脏病发作，钻石被偷的前一天就已经死了；我碰巧就是签署死亡证明的那个医生！”华生的证词会推翻警察局长的话。那么，三段论的逻辑性便会迫使福尔摩斯至少抛弃其中一个前提：也许小偷既不是男管家也不是园丁，而是其他人，或者说小偷根本就是男管家。

大多数普通探究和论证都与经验性事实相关。它们包含一些前提——即便我们承认其正确性——也不是必然正确的，因为会有例外。从这些前提中得出的结论“继承”了前提的不稳定性。若某个三段论的前提严格意义上不算正确，那么该三段论便不是严格的证据。因此，若是拥有充分甚至令人信服的理由来拒绝接受该三段论的结论，那么这应该会促使人们重新考虑这些前提——是的，重新考虑，但不一定要完全拒绝。人们可能会认可前提的例外情况，而不是简单地否认这些前提。

为了进行阐释，我们修改了“玛丽要写论文”的案例（鲁思·拜恩利用这个例子来证明演绎推论可以被“废止”了，正如我们在第一章中表述的那样）。

唐（Tang）和茱莉亚（Julia）是玛丽的室友。她们想知道玛丽是否能准时赶回来吃晚餐。

唐：嗯，如果玛丽要写论文的话，她会在图书馆待到很晚。

茱莉亚：实际上，她的确要写论文。我们不等她了吧！

因此，她们准备好了晚餐，摆好了餐具，打开了红酒，然后就在她们准备坐下用餐时，玛丽回来了。

茱莉亚：我们以为你要写论文，会在图书馆待到很晚。

玛丽：我不用写论文，如果图书馆还开着，我会待在那儿。

唐：要一杯红酒吗？

唐和茱莉亚最开始的陈述为假言三段论提供了两个前提，该假言三段论的结论是“玛丽会在图书馆待到很晚（因此不会及时赶回来吃晚餐）”。她俩最初都认可这个结论，但后来，当玛丽出现的时候，很显然又不得不推翻这一结论。不过，玛丽的解释显示她们的前提是可信的，且其推理也是明智的；问题仅仅是条件不那么正常了。

像唐这样的陈述“如果玛丽要写论文的话，她会在图书馆待到很晚”，并不是要表达或被理解为要表达某个必然真理或毋庸置疑的经验事实。像日常生活中大多数此类陈述一样，它在正常的条件下展现了较高的可能性——但是，条件不总是正常的。唐没有必要限定其陈述，说些“很可能”或者“在正常条件下”之类的话；不管怎样，茱莉亚都会这样

理解。玛丽出人意料地出现表明该三段论的两个前提不是让人认可其结论的充分条件——不是因为其中一个是错误的，而是因为其中一个可能会有例外，我们日常生活中的大多数归纳也是如此。

所以，事实是，使你否认该三段论结论的理由甚至不会迫使你否认任何一个前提；它们只会让你在其中至少一个前提中，更加意识到正常条件不变（也称“其他条件不变”）这一特性。

对此做何解释？人们可能会否认“如果”“且”“或”等连接词源于它们的传统逻辑意义。或许其意义基于非单调逻辑，或许是概率性的。这些语义上的解决方式表现为修改我们对这些词义的理解。然而，还存在另一种方式。哲学家保罗·格莱斯赞成用语用方式作为替代，该方式利用逻辑连接词（实际上一般指的是词语和表达）来解释发生的事情，不是过多关注词义，而是关注说话者在使用这些词时的言中之意。他认为可以保留逻辑连接词的传统词义，并认为可以用语用术语来解释显而易见的反例。（当然，语义和语用方式并非不能兼容；相反，它们应该被融合到一个完整的表述中。）注

我们在此大致简评了连接词和三段论的语用意义。格莱斯的深刻见解在现代语用学，特别是关联理论中得到了发展。关联理论的基本观点是，词和句子的语言意义并不是要将说话者的言中之意进行编码，而只是表示它——以一种更准确的方式来表示它，但留有解释的余地。

一游客问某个巴黎人：

游客：埃菲尔铁塔还有多远？

巴黎人：就在附近。

就巴黎人的回答而言，如果问路的游客是步行，“附近”这个词可能表达“步行距离很近”的意思；而如果游客问路时是在车里，则表达“行

驶距离很近”的意思。“附近”这个词当然有语义，但是这一意义与说话者想表达的那个意义并非完全等同；联系上下文语境，可以从这一语义推断说话者的言中之意。

“附近”这类词意义很模糊，只有放在上下文语境中才能更准确。那么“笔直”这类有准确意义的词又如何呢？这些都只是作为推测说话人意思的指示器。

游客［站在奥赛码头（Quai d'Orsay）的美式大教堂前］：去埃菲尔铁塔怎么走？

巴黎人：笔直向前走，你会看到的。

实际上奥赛码头并不是一条笔直的街。它沿着塞纳河的曲线修建。如果游客按照字面上的建议，笔直向前走，最终会掉进河里。不过，在上下文中，“笔直”传达的意思是应该待在同一条街道上，即使路径是弯曲的而非笔直的。“笔直”这个词在此用得比较随意。所以我们可以使用词语来指示比语言学上的意义编码更狭窄或更宽泛的意义。[注]

逻辑连接词（及数量词、情态动词）表现得就像一般词汇一样。它们没有为沟通者欲表达的意义编码，而仅仅是将其指示出来。拿“或者”举个例子。“或者”的意义是这样的，如果两个析取项（P、Q）中有一个是正确的，那么“P或者Q”形式的陈述就是正确的。所以，举个例子，如果园丁偷了钻石，那么福尔摩斯说的“男管家或者园丁偷了钻石”就是对的。但是，正如我们指出的那样，福尔摩斯通常会被认为不仅是想直接表达字面意思，还想暗示其有某些理由来断言该析取命题，而不仅仅是知道析取项中的某一个（在这个例子中，即园丁偷了钻石）是正确的。通常，相比每一个析取项，“P或者Q”陈述表达了对析取命题本身更大的信心。这样，在大多数普通的语境中，“或者”这个词便传达了比其编码的逻辑意义更多的意义。

“或者”也能用来传达少于其逻辑意义的内容。设想一下下面这段对话。

> 警察局长：男管家或者园丁偷了钻石。
>
> 福尔摩斯：我们确定吗？男管家或者园丁偷了钻石，或者房子里其他人偷了钻石，或者该村庄、郡里或国内的其他居民偷了钻石。

如果结果表明小偷甚至不是本国的居民，而是来自国外的游客，那么指出福尔摩斯的错误不会有太大意义。很明显他用了多重析取（“……或者……或者……或者……”）来表达一种隐性含义，即说话时带着“或者”的一般性表达：每个由“或者”连接的单独命题都是有疑虑的，他在没有承认整个析取命题的真实性时便这样做了。换句话说，他是在表达多于或少于其话语字面意思的内容。

其他像“和”“如果”这类逻辑连接词（以及“一些”“全部”这类数量词）不能（或不仅）用来传达其字面意思，而是用来传达从上下文中推断出来的意义。例如，众所周知，“如果P，那么Q”论述可能旨在或理解为不仅要表达“P是Q的充分条件”（这与“如果”的字面意思相一致），而且要表达“P是Q的充分必要条件”（与“如果”和“只有……才”相对应）或者是表达“P是Q的必要条件”（与“只有……才”相对应）。当沟通者用这种间接的方式来使用“如果”时，他们便没有违背任何逻辑准则或语义准则。他们只是正常使用语言。事实是，他们可以用与“因此”“但是”这类对话标志词相同的方式来使用逻辑连接词，并以此表明，即便语言的字面意思没有暗示什么言外之意，也可以从上下文获知。

这可能会让人认为，逻辑关系的语言表达与算数关系（如“135欧元被27欧元整除为5等份”）的语言表达类似。不论它们是否用书写数字及特殊符号或词语来完成，算术操作都遵循构造和解释的严格法则。即便

算术是在口头上执行的，对其操作的解释也不是基于语用考量；数字词和“整除”及“相等”这类词都是用的字面意思。

不像口算根据其自身法则用词语来开展业务，论证不是借用口头工具的逻辑业务，它完美地适应普通语言沟通的构造。但绝不能离开惯常的表达性及解释性语言实践。

有逻辑连接词（或其他逻辑策略）的陈述，即使这种陈述的序列或多或少与三段论相一致，都只是正常语用的一部分。它们被说话者用来表达一种意义，这种意义不能仅仅是被解码，还要旨在被进行语用学解读。不仅陈述所使用的词语，而且陈述的前提和结论被提出时所用的力度，都可以被解释。它们可能会被作为绝对的或不确定的断言，并受到隐性的“正常情况下的条件”的束缚。

当你辩论时，你一直在用正常方式使用语言，你的受众也没有避免使用其一直使用的同样的语用能力来解释你的陈述。在论证中，普通的表达及解释形式并未被所谓可以与算术法则相媲美的“推理法则”推翻。我们要接受算术法则，而且不能质疑它。但是，我们在推理法则的内容和存在形式上没有达成一致意见。有时我们学到的推理法则，要么是基础逻辑，要么是对所谓合理思维或合理论证靠不住的建议（如要避免的谬误清单，其本身就是有误的）。[注]

我们一直把注意力集中在论证上，但是“人们大脑中”的推理又如何呢？我们认为，这种个人推理正常使用了静默的内部语言。[注]内部语言的语用还没有被充分研究过。然而，鉴于许多内部语言都会演练或预测与他人的对话，所涉语用与公开演讲中的语用相比，很可能并无太大不同。我们自言自语时同公开演讲时一样，会随意地、修辞性地运用词语。我们的断言很可能被隐性的“很可能”或“在正常条件下”束缚或限制了。没有理由假定，我们在自己大脑中进行推理时，遵循了在公众论证中通常被忽视的逻辑法则。

难道所有显示普通推理没有受到传统逻辑法则支配的证据，都表明普通推理受到另一种逻辑或概率推论体系的支配吗？传统的演绎逻辑是“单调的”。这意味着，如果某一结论逻辑上遵循最初的某套前提，便同样也遵循任何包含最初那套前提的更大的一套前提。正如“玛丽要写论文”这个例子阐明的那样，人类的普通推理不是单调的。不仅在日常生活中，而且在科学、技术、医疗或法律推理中，结论通常是不确定的。鉴于新的考量（如玛丽最终及时赶上了晚餐），结论会被修改或被撤回。实际上，单调的推理对现实认知体系充其量只会发挥有限的作用。

在早期推理实验心理学中，普通推理的非单调特点在很大程度上被忽视或理想化了。而最近的许多推理方式却将其置于中心地位。一些学者，如认知科学家凯斯·斯坦宁（Keith Stenning）和逻辑学家米歇尔·范·拉姆巴根（Michiel van Lambalgen），旨在用一个“非单调逻辑”来替代传统逻辑，这会为人们的实际推理方式提供更好的洞见。其他人，如心理学家迈克·奥卡斯福特（Mike Oaksford）和尼克·查特（Nick Chater），则认为最好不要把推理视作逻辑的，而应视作概率的——更确切地说，是贝叶斯式的——思维形式。注

把标准逻辑替换为非单调逻辑或把所有逻辑替换为概率的计划，与传统理论共用一个基本假设：推论研究必须基于对优秀推论规范的一般性及正式解读。我们并未被说服。我们赞成推论过程进化的、模块主义的观点。每一个推论模块都旨在提供某种具体的认知利益，并且以低成本的高效方式这样做。在此视角下，对给定模块的研究便是把其特定步骤与其特定功能联系起来。特别是，理性模块的功能一般比优化知识及决策的功能更加具体。研究推论的一般规范，虽然从哲学或机械智能的角度来看很有趣，但不会告诉我们有关任何特定推论模块的太多知识。它只会告诉我们有关理性模块的极少数知识（正如运动的一般理论在理解蝙蝠飞行或蛇爬行方面用处有限一样）。

虽然我们不知道最近的非单调或概率理论是否会为充分理解推理打开大门，但无疑会认同批评家对传统逻辑的看法，即传统逻辑没能为人类推论提供貌似有理的规范。难道我们不应该也认同传统逻辑完全与推理研究无关的看法吗？好吧，不是这样的。让我们冒险为传统逻辑做个有限的辩护吧，对此，传统逻辑的捍卫者可能不会太开心。逻辑不仅能用作规范或程序，而且能用作启发式工具来阐明问题并给出答案。我们认为这便是逻辑在推理中发挥的主要作用。当然，这与标准观念背道而驰，标准观念认为逻辑的功能恰好能克服启发式思维的有限性。

如果我们解释三段论时可以有某些自由，而且根据上下文来推断，即使这些三段论是符合逻辑的，也不一定会迫使理性的推理者认可其结论，那么它们又有什么用呢？使用它们的意义到底何在？我们想要表明，在认可某个非瞬间直觉结论时，这些三段论通常突出理由；在否认某个瞬间直觉结论时，它们亦是如此。三段论的图解式说明（及大致的演绎关系）往往夸大了我们认为的对逻辑的依赖程度。特别是，它把纯粹的不连贯性（其中，两个及以上的观点会各自给出否认其他观点的理由）戏剧性地表现为明确的不一致或逻辑矛盾。正如扩大图片的轮廓能提高辨识度，正如在叙述中丢掉细节能帮助人们把握整个脉络，丢掉束缚、忽视例外有助于人们把注意力集中在会促使其认可或否认某个结论的理由上。不连贯性通常很难被察觉和反思。它们表现出的逻辑上的不一致使其成为理性模块的抨击对象。[注]

正如我们描述的，推理本身涉及对低阶直觉如何支撑某一结论的高阶直觉。的确，在原则上，高阶直觉可能仅与理由和结论的逻辑相关性有关，可能会忽视其内容的其他方面，但它们为何这样呢？总之，这不是事实。当就某个给定问题进行推理时，高阶直觉有关低阶直觉的各种性质，不论是不是“逻辑的”，都与其作为理由的价值相关。在推理中，高阶直觉是元认知性的，而非仅是“元逻辑性的”。推理基于丰富多样的对直觉的直觉。

我们借助大前提是“P或者Q”形式的析取三段论，来采用这种传统意义上会被表征的推理。就是说取决于其中的P和Q，即取决于从标准的逻辑学观点来看不应被考虑进去的内容性质——高阶直觉可能会向不同的方向发展。

你在推理所需的某本书的位置，并凭直觉知道这本书一定在书房或起居室。你在书房找过，但是到目前为止还没有找到。因此你决定去起居室找找。当然，若被要求解释为何现在要去起居室找，你可能会以三段论的形式给出明确的回答：“我认为书一定在书房或起居室，它不在书房，因此肯定在起居室。”这种三段论形式可能会使你的回答看起来有点逻辑。实际上，你的推理与该案例的细节更合拍。你在脑海中并没有排除可能遗漏了书房某个角落的可能性；你寻找过程中失败的次数越多，这一直觉变成推断出“书肯定在起居室”的直觉理由的概率就越大（不然的话，你最初认为“书一定在其中一个房间”的直觉便错了）。你的推理过程易受到“书不可能同时出现在两个房间”这一事实的影响，也易受到“两个房间都很杂乱无章，你可能没看到”这一事实的影响，也易受到“你最初关于书的可能位置的直觉只是基于不太确定的记忆”这一事实的影响。即使你把它表述为一个三段论，你还是希望受众仅将之视为你思维的图解描述，并希望他们使用自身更丰富的直觉来理解并评估你的理由。

现在举个不同的例子，这个例子中的图解式说明是相同的，只是相关的直觉截然不同而已。看到莫莉皱眉了，你凭直觉知道，她一定在为昨天发生的某件事而沮丧，或是担忧可能会发生某件事。到底是什么？她告诉你昨天发生的事情并没有使她沮丧，你猜测她的话属实，便因此推断出她很可能在担忧。你最初认为莫莉可能很沮丧或者担忧的直觉并不彻底排除她可能两者兼备的可能性。她说自己没有很沮丧时的真诚并不排除“她可能很沮丧，只是没有意识到而已”的可能性。那么，鉴于你支撑某个稍不确定的结论的直觉理由，你往往会认为她在担忧而非沮丧，或至少担忧多过沮丧。假设你现在告诉你俩的朋友雷蒙

（Ramon）：“莫莉忧心忡忡！”他回答：“我认为她仅仅是因为昨天发生的事情沮丧而已。”然后你可能会以三段论的形式进行论证：“看着莫莉的脸，你看不出她是沮丧还是担忧；不过，她说了没有因为昨天发生的事或其他事情而沮丧，而且她显然很真诚；因此她一定在担忧。”这会再次图解式地说明你的推理，但这样做，会挑战雷蒙的观念，并给出让其修改观念的理由，或让其仅推断出莫莉在担忧，或以一种更细致入微的方式来看待事情，正如你自己所做的那样。

特别是在论证式推理中，推理者运用逻辑关系会对其受众发挥启发式作用。它会帮助受众去检查、丰富或修改其信念，或反过来用论据维护自己。在某种程度上，多亏了其逻辑形式，即使论证不总是令人信服，至少也极具挑战性。

然而，一般而言，推理不是运用逻辑（或任何相似的正式体系）来获得结论。那么，推理方式又是什么呢（如果有的话）？

有推理方式吗

正如目前为止我们描述的那样，推理是颇受限制的。当人们试图去说服他人或被他人试图说服时都会用到推理。

单独的推理似乎会在以下几种情况下出现：期待与他人进行辩论或重新进行换汤不换药的辩论时；当人们发现自己所持观念自相矛盾时，或者在与自己进行辩论时。与正当理由一样，论据的产生也是通过逆向推论，即从所赞成的结论推出支撑该结论的理由。

很显然，在这一图景中，一定遗漏了某些事情：人们通常不是因为想法的碰撞（与他人或自己），而是因为自问自答的问题刺激才进行推理的。此外，难道这种个人探究式推理的意义不在于找到某人自身问题的正确答案，而在于证实某个已赞成的结论吗？

我们确实可以靠自己来解决一些问题，不带任何偏见，不需要有对这个或那个答案的预感，我们可以通过从前提到结论的顺向推理来解决这些问题。不过，这些往往是特殊的问题，要通过特别的推理方式予以解决，而不是用每个人的日常推理倾向。我们可以在用来娱乐、教学或测试的游戏或拼图中找到这类问题最简单的例子。

拿数独（Sudoku）游戏举个例子（如图9.1所示）。该游戏是在一个由81个小格子构成的正方形网格中进行的，其中一些小格子在游戏开始时就已经填了数字，其他的则空着。

5	3			7				
6			1	9	5			
	9	8					6	
8				6				3
4			8		3			1
7				2				6
	6						8	
			4	1	9			5
				8			7	9

图9.1 数独

该游戏的任务是找到每一个空格子该填的数字，要知道1到9中每一个数字只能在每一列、每一行、每一个用粗线标识出来的3×3的格子里出现一次。

数独玩家用公平的方式完成了游戏任务。他们没有事先的预感，也不拿特别的解决方式来冒险。他们知道，在每个网格中，对每个小格子来说，有且仅有一种正确的解决方案，玩家的任务就是填满所有的空格。这样就很明确了。只是玩家继续游戏的方式不那么明确罢了。

在最简单的情况下，通过简单的排除法，直到只剩下一种可能性，便可以找到能放进给定小格子的数字。例如，中间的小格子（图中灰色的）里应该填哪一个数字呢？看着这个网格，你排除了数字1（因为它在包含灰色格子的行、列里都已出现过）、数字2和数字3（因为它们在3×3的格子里已出现过），除了5之外，其他数字都可以这样被排除，因此中间小格子里一定是填5。九宫格是夏洛克·福尔摩斯著名格言的完美阐释，其格言是："当你排除了所有不可能的情况后，剩下的，无论多么不可思议，都一定是真理。"（"像夏洛克·福尔摩斯一样思考"的用处非常有限。）

我们刚刚阐释的排除法可以帮助游戏玩家找到一些缺失的数字，却不能帮助他们填满整个网格。要彻底完成某个数独拼图，即使是比较简单的数独拼图，玩家也必须使用更复杂的方法。不过，对心理学家而言，真正的挑战与其说是理解玩家如何应用其学会的方法，不如说是理解他们某些人如何至少独自找到有用的方法。[注]毕竟，找到这些方法的推理比你知道该应用哪种方法来完成拼图的推理更令人钦佩。

普通的数独玩家要如何找到新方法呢？我们认为，一般而言，当人们进行推理时，他们从赞成某一给定结论的直觉偏见或预感出发，然后通过逆向推论寻找支撑该结论的理由。我们认为，尽管情况并非玩家"运用"方法来破解数独拼图，然而却是玩家起初（这时他们不向他人

学习）就“找到”这些特殊的方法。仅仅通过理解游戏的规则，我们就能凭直觉运用上述的简单排除法。玩家通过实践熟悉了数独网格各种难以察觉的规律，并培养了对愈加烦琐的方式的直觉洞见，对这些规律和洞见的成功应用便证实了其实用性。

找到完成数独拼图的有效方法，很可能是基本的、有偏见的推理的又一实例：寻找支撑最初直觉的理由——这种情况下的直觉不是把数字放在给定小格子中的直觉，而是一种高阶直觉，即让玩家缩小每个小格子中数字可能性范围的考量。如果我们正确，便意味着人们没有使用一般的、高阶的明确方式来找到各种所需的更具体、更明确的方式来完成数独拼图。

数独拼图是人们通过实践及明确的方式可以解决的问题。这可以延伸到各个领域的推理问题吗？人们使用正确的方法可以有效回答各类问题吗？答案是否定的。

是的，人类有能力运用他们学会的方法或独自找到的方法来做各种各样的事情：完成数独拼图、建造乐高城堡、在按照字母顺序编排的字典中查单词、从自动售货机买东西、烘焙蛋糕、把罗马数字转换成阿拉伯数字、学习踢踏舞的基本步法、找到算术或几何问题的答案或使用数据库来回答关于判例、预期寿命或股票收益率的疑问。这种理解并运用按部就班的方法的能力是人类心理学非常重要的方面，在文化技能（包括解决问题的专门技能）的发展和传播中发挥着重要作用。心理学家基思·斯坦诺维奇认为人们获得并运用这些方法（或“决策支持系统”）的从容度与个人理性差异相关联。[注]尽管如此，推理不在于应用这些方法，总之，推理不需要这些方法。正如军队行军不是最典型的步行，对这些方法的运用也不是最典型的推理。

人们在日常生活或在追逐自身长远利益时遇到的大多数问题，无论怎样，都不能在下列指示中得到回答。每年，新出版的心理咨询书都会

为你在事业、爱情、友情和游戏中如何进行更好的推理提供指导，并承诺会比去年的书做得更好，其实去年的书也这样承诺过。虽然它们有时会给出某些明智的建议，但这类书的通病便是不能兑现承诺。要得到这种理想的结果，我们需要理解我们所处世界的许多方面，需要知道在给定的情境下该考虑什么、该忽略什么；需要识别并解决沿途出现的各种问题；并无充分指示来让人对大多数现实问题进行有效推理。

数独拼图十分明确的定义使其与大多数我们可能会推理的问题相区别。上下文再宽泛也没有用，证据就摆在那儿，有且只有一个正确的答案，而且任何不正确的答案都会导致明显的不一致。在心理学中，研究这类有明确的、可辨识的答案的问题及某些能有效找到其答案的方法时，通常打着“解决问题”的标签而非“推理”的标签。在我们还不知道正确方法时，解决这类问题不仅需要某些推理，还需要试错修补以及与任务相关的特定形式的洞见。

从这个意义上讲，许多方面都要用到问题解决能力，如许多游戏中及数学、逻辑、编程或技术上的严肃问题。科学家、技师和外行人都已发现并研制了可以用此方式解决问题的一套方案。这套方案中的某些方法因其实用性、科学关联性、知性的优雅或作为休闲活动的吸引力已成为成功的文化案例。

对推理心理学家而言，那些可以用明确有效的方式来解决的问题和拼图，特别是那些实验者专门设计的问题——如沃森选择任务——似乎能为实证研究提供最理想的材料。他们总是认为，其中某些问题不仅实验起来容易又有趣，还会为基本推理机制提供关键证据。不过，这种狭义界定的任务只是推理帮我们应对的各色挑战中的特殊案例。“解决问题”需要用到一些推理，但狭义上说，大多数的推理甚至都不会与“解决问题”相仿。

人们在推理道德、社会、政治或哲学问题时，几乎不会达成普遍的一致。可能他们都认为，对这些普遍的问题肯定有一个正确的答案，一

个每位有能力的推理者都会认可的答案。他们可能认为，那些不认同他们的人，如果不是存心欺骗，便是不理性的。但是认为"只有你本人及那些认同你的人是理性的"，这一想法的理性何在？更加貌似可信的结论便是，推理不论多么合理、多么理性，都不能确保想法的趋同。

是的，科学家的确会达到某种重大且渐增的趋同程度。这在某种程度上可能要归功于许多被认真研究出来的方法，它们在指导和评价科学研究中发挥了重要作用。然而，某些发现或突破却没有得到任何指示。从科学家自己的解释中可以判断出，重大的突破源自某些预感，这些预感在搜寻的过程中得到发展及微调，通过逆向推论，在证实论据和证据的同时，阻止并逐渐削弱竞争者的反对言论。

可以说，科学是汇集人类血汗的领域，其中理性和合理推理都是最有价值的（我们在第十八章中会详细论述）。不过，推理不会促使科学家就最好的理论自发地实现趋同，反而会促使其详细阐释并拼命捍卫各自不相容的、相互抵触的理论。推理也会帮助他们——这次伴随着更高的趋同程度，即便仍不完美——去评估这些相互抵触的理论，并对该竞争中当前赢家的理论达成某种程度的暂时性一致。在很大程度上，赢得这场竞争很可能是认知运气问题——寄希望于最初更强的预感，且能找到更强有力的证据及论据来支撑该预感——而不只是产生新想法时较合理的推理。虽然科学实践表明，各种具体方案可以与具体推理任务相关联，但其实并没有一般性的推理方法。

推理研究受到规范性目标及描述性期望的支配：规范性目标是发现并明确能产生合理推理的普遍方法，描述性期望是人类实际推理时会受到近似方法的指导。到目前为止，该目标和期望都让人大失所望。这并非偶然。对此有一个原则上的解释。我们认为，直觉推论程序是无意识的、机会主义的且多变的。认为"直觉可能表现为遵循某个一般性方法"没有太大意义。

我们对支撑某一直觉结论的理由的高阶直觉把许多低阶直觉的性质

考虑进去了，这些性质可能在领域间的变化很大。这并不是不合理的推理，更不用说荒谬的了。在以论证形式摆出理由说服他人时，我们有时会求助这些理由本身具有细微差别的高阶直觉，有时会选择图解式说明，其实我们通常两者都会涉及一点。我们所应用的、根植于日常用语中的图解式说明，激励了人类历史上一项伟大智力成就的诞生：逻辑科学的发展。然而，逻辑既没有告诉我们怎样推理，也没有告诉我们应该怎样推理。不存在我们可以而且应该遵循的一般性推理方法，我们要么靠自己，要么与他人进行交流。

---

1. Mercier & Sperber 2009; Sperber 1997.
2. 从广义上看，把“直觉的”当作某人有直觉论据的反思结论很常见。用此举来延伸意义并无妨碍，所以我们一般不反对。不过，在本书中，我们在论述直觉理由与其反思结论之间的联系时，所使用的“直觉力”和“直觉的”绝对没有延伸意义。
3. Whitehead & Russell 1912.
4. 推理有时也会依靠图片和手势，因为它们可能要比话语更有表现力，但是这些表现力通常只会用在口头上，用以支撑口头的结论。可参见：Stenning & Oberlander 1995。
5. 也许有人会说，元表征所表征的优先能力让物种在进化过程中产生了语言，让个人在发展过程中习得了语言（Origgi & Sperber 2000; Scott-Phillips 2014; Sperber & Wilson 2002）。这一观点并没有得到普遍认同。反对观点认为，有语言才有了元表征，因此在进化及发展过程中都是先于元表征的，可参见：Dennett 2000; De Villiers 2000。
6. Blakemore 2002; Ducrot 1984; Ducrot & Anscombre 1983; Hall 2007; Wilson 2016.
7. 哈曼（Harman, 1986）对此进行了强调。
8. 欲从对话的角度了解归谬法（reductio ad absurdum），可参见：Dutilh Novaes 2016。
9. 可参见：Evans & Over 2004。
10. 欲了解详细论述，可参见：Carston 2002; Wilson & Sperber 2002。
11. 可参见：第十二章及Mercier et al.即将出版的书籍。
12. Carruthers 2006.
13. 可参见：Oaksford & Chater 2007; Stenning & Van Lambalgen 2012。顺便提一下，这两本书都建议重新解释拜恩的“抑制效应”，也就是指可以通过添加逻辑上并不相关但实际相关的前提来抑制演绎推理，正如我们在第一章中看到的。
14. 当然，严格意义上且直白地来说，我们并不是在否认确有问题可以通过三段论法或

其他演绎形式解决。正如我们在第二章论述的沃森四卡选择任务的例子中阐明的那样，即便问题可以通过演绎的方式解决，人们通常都没能做到。尽管如此，特别是当人们辩论这类问题时，最终都会以标准化形式援引逻辑性考量。

15. 这一问题在Louis Lee，Goodwin & Johnson-Laird 2008的书中提了出来。

16. Stanovich 2012.

第十章
理性：为了什么

看一看图10.1中出现的一件家具：几乎算不上一把设计精美的椅子，不是吗？座部太低了，让人几乎得蹲坐在上面。椅背又太高了；顶部的横杠是水平的厚板，让人不能舒舒服服地倚靠在上面。怎么会这样呢？实际上，这不是一把椅子，而是一个教堂跪垫。这个高度正适合跪下。“背面”实际上是前面，顶部的横杠在这儿不是用来倚靠的，而是祈祷时用来放置手的。一旦人们认识到了这个手工艺品的真正功能，之前看起来的瑕疵就变成了精心设计的特点。

就像图10.1中的这个物件，理性似乎也有一个很明显的功能：帮助个体独自增长见识并做出更好的决定。毕竟，如果运用理性不能帮助人们获得更好的信念、做出更好的选择，那么，理性作用何在？然而，就像被当作椅子的跪垫那样，推理在表现该功能时也十分糟糕。

我们已仔细观察了众多推理失败的案例，且在第五部分会进一步论述，但对如何推理的实证研究结果表明，我们在用一种带有偏见且懒惰的方式进行推理，而且经常没能解决简单但不熟悉的问题。就像许多心理学家所做的那样，止步不前并简单地认定人类理性只是设计得很糟糕而已很容易。

也有另一种观点，认为人们可能弄错了推理的真正功能。在现代心理学中，推理可能就像被误认作椅子的跪垫。在第七章至第九章中，我们已证明，人们通常在追逐社会互动目标时用到理由，特别是在为自身辩护及说服他人时。在此，我们采用进化理论来证明，理性的这些社会

功用并非把提升个人认知作为其主要功能的认知机制的不良反应或次要功能，相反，理性的主要功能是社会性的。为什么要诉诸进化理论呢？因为这是唯一可以解释以下事实的方式：生物复杂的遗传特点往往会产生一些有益的影响。如果不从进化视角来看的话，要弄清人类理性及与之相关的任何事情到底为何会起作用，会变得非常不易。

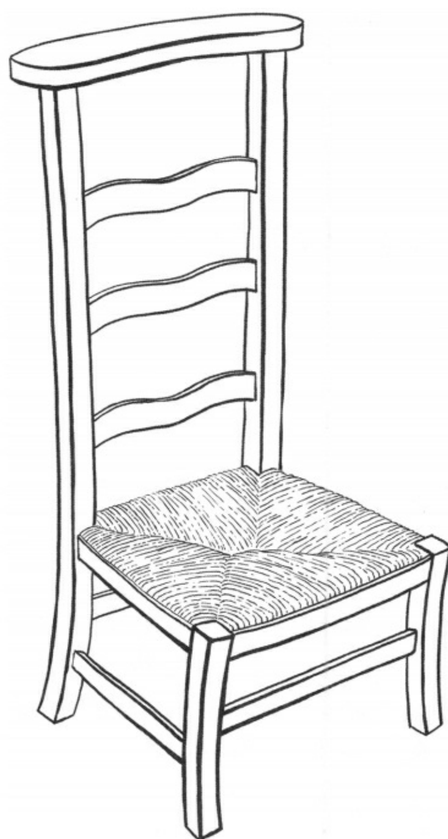


图10.1 设计拙劣的椅子？

达尔文之前及之后的功能

人类理性有其功能，这通常被视为理所当然，我们不想去质疑这一观点，而想在新的基础上对其进行发展。我们先要弄清楚功能的概念：对这一概念的肤浅理解通常会损害对理性进化的论述。

功能是什么？人们要怎么弄清像椅子这种手工艺品、花香这种自然

属性、翅膀这种身体组成部分或人类理性这种心理能力的功能呢？功能是什么？我们如何识别某个特定功能？这两个问题紧密相关。对手工艺品，我们通常很容易给出大致的回答。手工艺品的功能就是指它的用途。谁决定手工艺品的用途呢？设计它的人。他们决定椅子是用来坐、跪，抑或是其他用途。

我们并不能直接把某一功能归于某一生物特征。没有人做翅膀用来飞，或做心脏用来输送血液。然而，又有谁会怀疑翅膀是用来飞的呢？谁又会对威廉·哈维（William Harvey）发现的心脏功能提出疑问呢？

在达尔文之前，把某项功能归于动物或植物的某一特征，只需要回答下列问题：它有什么用处呢？找到某一生物特征的好作用，并假定这一好作用便是其用处。有什么能比飞翔更有用呢？所以，翅膀的功能便是飞翔。这种功能的常识性概念并不是没用的，而是在达尔文之前的科学中，就发挥着重要的作用（如在哈维对心脏功能的发现中），而且其在现代科学中至少仍发挥着启发式作用。这一与历史无关的功能概念提出的且没能解决的问题是：为什么我们一开始要期待生物特征具备有用的功能呢？我们为什么不期待其他自然物体，如石头、行星和亚原子粒子有任何功能？这些物体没有功能。长期以来唯一的回答便是，上帝这个造物主做出了用来飞的翅膀、用来输送血液的心脏等，这一回答用深不可测的谜代替了一系列有趣的问题。

达尔文的自然选择理论为清晰且有用的功能概念提供了依据，这一功能概念基于对生物器官及其他生物特征为何具备其所具备功能的真实解释。

任何生物的遗传特点都有许多影响。其中大多数影响不会促成成功繁殖。例如，花香可能会使附近的某些动物头晕目眩，但这种花香既不会伤害植株，也不会使植株受益。某一特性的某些影响可能碰巧会增加成功繁殖的概率：花香会吸引昆虫，并由昆虫来充当传粉的媒介。但是，该特性的其他影响实际上会抵消成功繁殖的概率：花香可能会招来

食叶昆虫，并诱使它们吃掉这朵花，从而减少传播的花粉数量。

当某一遗传特性大体上对促成其携带者成功繁殖利大于弊时，这一特性便很可能会被选定并世代相传。被选定的特性中至少有一个影响会促成其携带者成功繁殖，从而有利于该特性自身的世代遗传。例如，野花的香味被选定便是因为该香味通过吸引传粉媒介促成了其植株的成功繁殖。当生物学家谈及选定特性的功能时，指的就是这样一种有利的影响。

被选定的特性可能不止一个功能，而且这些功能在进化的过程中可能会发生改变。例如，企鹅是鸟类的后代，它们曾经用翅膀来飞，可现在其翅膀的功能是游泳。在它们进化的某一时刻，这些翅膀可能同时具备两个功能，正如某些海鸟（如海雀）的翅膀那样。

有些特性能够世代遗传，多亏了其本身具备的功能，而这些特性就是适应的结果。但不是所有生物特性都是适应的结果。某些特性源自生物发展的物理或化学约束条件。例如，“所有生物体都含碳”这个事实便是生物本身基本的化学性质（至少在泥土中找到了）。而且，并非所有的适应特性都是适应性的。许多适应特性都有消极的影响。心脏有规律地收缩就是一种适应特性。这种收缩有让血液流至全身的功能。但是，由这些收缩造成的声响并不是适应性的，而是负面影响。

理性的功能——从亚里士多德到20世纪的心理学家

就像有关翅膀功能的观念一样，有关理性功能的观念在达尔文及功能的科学性概念出现之前便已出现。即便在达尔文之后，许多理性进化思维与达尔文之前的观念也有同样的认识，即认为理性的功能就像翅膀的功能那样，相当明显。理性是个人获得更高深知识、做出更明智决定的手段，理性通过发挥这一智力功能，使人类优于其他动物。

达尔文本人也表达了与此相同的大致观点，他在《人类的起源》（*The Descent of Man*）一书中写道："在人类心智的所有能力中，我料想，人们会承认理性是处于顶端的。"[注]不过，他用进化论扭转了这一观念。他认为，人类运用智力功能，"拥有了使自身习性适应新生活环境的强大力量。他们发明武器、工具和各种计谋来获取食物并保护自己。当他们迁徙到寒冷的地方时，会用布裹身、搭小棚屋并钻木取火；用火煮食物——因为不煮熟的话，食物难以消化"。因为这些智力功能是多种多样的，而且往往能够遗传，"它们会通过自然选择进行完善或者优化……在社会最蒙昧的时代，那些最睿智的个体……会养育最多的子孙后代"。[注]

达尔文眼中的理性功能的影响和亚里士多德的观点大致相同。不过，达尔文的标新立异之处在于运用这些影响解释了理性为何要进化。而且因为达尔文对理性功能的断言更精细，所以其断言更容易受到科学挑战。若理性的进化是为了帮助人们思考自身，那么依据认知利益和心理成本，这种进化确实应该培养更好的思维，如果做不到的话——正如现代推理心理学中的许多内容表明的那样——那么我们真的就面临严重的问题了。

你可能会认为，那些找到在他们看来是人类理性重大瑕疵的心理学家，一定获悉并探讨了这对理性功能的后达尔文主义解读所带来的挑战。事实上，他们并没有。直到20世纪90年代，推理心理学文献都几乎没有提过进化论、自然选择及生物功能。心理学家只是理所当然地认为理性的功能——其中"功能"是常识意义上的理解而非生物概念——是提高个人认知水平。然后他们就推断出，理性并没有发挥其功能及所假定的功能。

甚至还不如说，心理学家把推理中的瑕疵当作反对进化理论的证据。进化机制只有很好地发挥功能，才能被选择。既然推理存在这些缺陷，便不适合探求知识及形成明智决定。因此，许多人断言，推理案例

显示了进化理论与人类心理的不相关性。[注]当然，心理学家对这一问题的常识性观念不该不加批评地认可，而是应该重新思考理性的功能，但他们没有这样做。

虽然心理进化理论的基础是由达尔文本人奠定的，而且历年来也出现了一些相关研究，特别是对动物行为的研究，但在心理学家勒达·科斯米德斯（Leda Cosmides）及人类学家约翰·图比（John Tooby）于20世纪80年代提出宏大的“进化心理学”计划后，对人类认知机制，特别是推理进化的一般的、系统的论述才得以开启。[注]

“自然选择塑造了人类推理的方式？”

从一开始，人类推理就已成为进化心理学研究的主要话题及反对此理论的学者论战的主要着眼点。当勒达·科斯米德斯于1989年发表了推理心理学史上最具争议性的文章后，论战就爆发了。该文章标题为“社会沟通逻辑：自然选择塑造了人类推理的方式？”（The Logic of Social Exchange：Has Natural Selection Shaped How Humans Reason?）[注]（因此便成为这一节的标题）。

在科斯米德斯的文章中，有两个主要的理论性断言。她认为，从进化视角来看，人们应该预计到，推理机制在处理远古环境中经常发生的特定问题时便已进化。专门进化来解决问题的适应性能力比任何解决问题的一般性能力表现得更好。实际上，我们还不清楚解决问题的一般性能力的基本成分（准确地说，如果其成分不是许多更专业化机制的复杂表达的话）。如果一般的推理能力能够进化，就可以帮助解决通过专门机制没能妥善解决的残留问题。科斯米德斯认为，在没有进化意义的任务上（如那些通常在推理实验中分配的任务），人们应该预计，推理的效率比专门进化的机制解决问题的效率要低得多。

科斯米德斯的第二个主要断言是，在许多最可能进化的认知性适应能力中，应该有旨在解决合作中主要问题的推论机制。合作是两个及以上个体间的互动，在合作中，每一个人都会蒙受损失或获得收益。比如，每个王牌猎人都会花时间并冒风险来帮群体捕到猎物。假设利益比成本大得多，而且平均分配，那么对于所有合作者而言，合作便是有利的。然而，在许多情况下，每个合作者都希望不付出全部成本便可瓜分合作利益，即通过欺骗或搭便车，获得更多利益。那些逃避风险却与他人分享同等份额肉的猎人便处于有利地位。当然，如果猎人都不去冒险，就会放跑猎物。广泛的搭便车行为很可能导致合作失败。要使合作有利可图而且得以持续，就必须识别欺骗者，控制他们或把他们排除在外。

科斯米德斯和图比认为，识别欺骗者应该被视为人类进化的主要问题，这促成了一个能计算合作者权利及义务的专门模块，人类以此为发现骗子的依据。为检验该假设，他们十分依赖沃森选择任务，该任务我们在第二章中已谈及，而我们认为沃森选择任务并不足以检验这一假设。[注]不过，我们的目的不是评估科斯米德斯的论据，其论据表明存在用来找出骗子的进化了的能力（我们认为该论据强而有力），也不是评估其实验证据（我们认为该实验证据较无力）。我们的目的是以她的理论为灵感及对照来源，以提出另一个不那么彻底但更宏大的进化假设，有关人类理性的一般性功能，特别是推理的功能。

像科斯米德斯和图比那样，我们认为进化的心智主要是由模块机制的组合构成的。模块是专门化的——每一个模块都帮助解决某特定类型的问题或利用某特定类型的机会，这在某种程度上有助于生物适应。另外，传统观点认为，推理旨在解决各类问题，并找到利用各种机会的方式。传统推理不是专门性的而是一般性的，不是狭小的模块而是广博的能力。

如果推理任务太宽泛了，以至于不能成为模块任务，似乎只能得出

两个结论：一是人类心智中主要的、最特别的成分不是模块性的；二是传统意义上的推理根本不存在。科斯米德斯和图比赞成第二个结论，认为不同于人们长期持有的哲学及心理学教条，推理通常并不存在。[注]其他从事心智进化研究的研究者，如心理学家塞西莉亚·海斯（Cecilia Heyes）和哲学家金·斯特瑞尼（Kim Sterelny）则赞成第一种结论，认为领域一般性推理的存在便证明人类心智的模块性其实比科斯米德斯及图比等人声称的要低得多。[注]

这是仅有的两种可能吗？不是的。正如我们已（在第六章中）展现的那样，元表征性模块处理特殊类型的客体——表征，并关注这些表征特有的性质。尽管如此，这种模块还是会间接提供与其处理的表征主题相关的信息。例如，“读心”不仅告知我们他人的想法，还告知我们这些想法涉及的所有事物及事件。我们认为推理产生于元表征性模块，该模块的特定范畴便是理由及其所支撑的结论间的关系。不过，这些理由及结论本身可以与任何话题相关。因此，对理由-结论关系的推论可以间接推出各个领域的结论，间接提供某种虚拟领域普适性。

所以，为回答科斯米德斯的文章副标题及这节标题提出的问题，我们一致认为，自然选择塑造了人类做出各种推论的方式，还催生了一系列专门的推论模块。我们补充道，其中一种便是理性模块。理性模块产生的正当理由及论据本身可能隐藏了与各领域知识及行为相关的推论。然而，这种虚拟领域普适性不会使理性及基本人类心智组织的模块性变低。

理性模块发挥什么作用呢？我们已摒弃了唯理智观念，即理性进化来帮助个体做出更好的推论、获得更丰富的知识并做出更好的决定。相反，我们赞成互动推理理论。我们认为，理性的进化是为了回答社会互动问题，而非独立思考所面临的问题。理性发挥了两个主要功能：一是通过拿出正当理由帮助解决主要合作问题；二是通过拿出论据来帮助解决主要沟通问题。（我们先前的研究工作一直集中于用互动理论推理并

发展“论证推理理论”。）注

协作面临的挑战及理性的辩护功能

人类超凡的协作不仅体现在规模上，还体现在所采用的各种形式上。其他动物可能在其全部行为模式中很少或完全没有为即兴创作留有余地，也很少有协作性互动。协作者间的相互期待很大程度上已预先确定了。当某个协作者因无能或背叛而无法满足这些期待时，协作便很可能失败。但其实人们不怎么依赖预先确定的期待。人们在利用现存协作形式时有很大的灵活性，且经常会根据需要随时调整。他们通常也会尝试新的协作形式。但只有投入大量能确保有效协作的认知资源，人类合作的这种灵活性及创造力才会变得非常有利。

人们要实现其各种协作形式所要求的更紧密的协作，需要让相互期待不断更新以保持可信度。例如，部队或运动队中的成员调整甚至重新确定其在不同场合的策略及相互作用。此外，人们必须相互依赖，不管是在事情按预期进行时，还是在事情未按预期进行时。因此，朋友、夫妻、同事及生意伙伴一定会相互依赖，即使这种依赖表现为不同形式。

在理解共同目标及每个个体必须发挥的作用时，即使是有能力的协作者，也会存在理解差异。利益分歧不会造成背叛或欺骗，相反，人们通常会基于要一致促成的公正来辨识并处理这些分歧；不过，理解差异也会导致人们的相互期待不能实现。

对人类而言，知道彼此在期待什么是关键的认知挑战。人类是如何应对这一挑战的？人类要如何成功形成即便说不上十全十美但至少恰当的相互期待呢？最常见的办法是援引两个机制：一是社会学层面的规范；二是心理学层面对他人心理状态的理解。

一方面，从社群层面来看，有各种共同规范要遵守，包括道德的、法律的、宗教的、咨询的、技术的等。这些规范有的被明文编进了法律法规，有的被强制执行，也有的并未如此。它们为各种社会互动定下了规矩。(注)

某些规范的直接目的在于保证某种对所有当事人都有利的协作。例如，若没有交通规则的话，开车出行会异常危险。然而，非常周密的协作法则往往是高度专业化的。例如，交通规则所规范的这种互动与其他任何互动都不一样；司机每时每刻都有相对狭小的、界定明确的选择范围。如果他们不协调彼此的决定，生命便危在旦夕。有效协作符合所有司机的利益。

另一方面，法律、道德或宗教规范，在许多方面都有利于各种社会协作，但是确保有效协作不是这些规范的唯一目的，也不必成为主要目的。这些规范有待进一步阐释，有许多相关问题还未涵盖在内。譬如，在某些情况下，晚宴上不迟到可能会变成失礼的行为。而且，人们在遵循这种规范时，利益并不总是相同的。例如，某些民选官员再次竞选时，会公然怂恿所有选民都按规定给其投票，对有些不想再选他们的选民而言，按规定投票有点困难。

为了实现有效的协作，社会中多种多样的规范是不是在以十分巧妙的方式表明人们的相互期待时限制了人们的选择？当然不是！许多法律、道德及宗教规范就告诉夫妻可以对彼此期待什么。但是，要实现日常协作所需的相互期待，比起理解“应遵守社会强制执行的某套规范”，夫妻双方需要更多地理解对方。

在许多社会互动形式中，一定程度的创造力和即兴创作是社会能容忍且需要的。例如，你要办一个派对，在思考该邀请谁时，传统惯例或道德标准会建议你邀请那些曾邀请你的人。因此，鉴于奥尔加（Olga）上个月邀请了你，也许你应该邀请她。对于邀请其他客人，互惠规范就

无能为力了，但仍有协作问题要处理。例如，如果你邀请迪亚戈（Diego），最好不要邀请露丝（Ruth）。如果你邀请露丝，她会让你邀请超（Chao）。

现在许多微不足道但相互关联的决定，决定了你今后与他人的互动。所以在处理规范无法处理的协作问题时，什么会对你有指导作用呢？标准答案是“读心”。你必须理解奥尔加、迪亚戈、露丝及其他客人的心理状态，以预料其反应，最终确保派对成功举行。这样，你会改善与他人今后的关系，而不是任其发展。

就像规范一样，“读心”在协作中也发挥了至关重要的作用。但是还远远不够完整。明显缺失的一点是，个体在推断彼此可以期待什么时，不只是基于对他人心智及共同遵守的规范的解读。人们仔细审查、讨论及协商对彼此的期待。许多与互动方式相关的决定本身就是通过互动做出的。

流言为可以对第三方期待什么提供了丰富的证据。不过，个体不需要成为流言的被动接受对象。人们可以参与到正在进行的、与自身相关的对话中，为自己的观点和决定做解释和辩护。而且，这样做的话，在某种程度上，会维护其自身声誉。有些官员试图封锁某些选票（因为这些选票不太可能重选他们），称此举只是为了防止选票造假，以此为自身辩护。派对上，有的客人来得太晚，会通过解释如何被耽搁来为自身辩护。你可以告诉迪亚戈，因为露丝之前邀请了你，所以你不得不邀请她；你可以告诉露丝，她当然可以带上超，因为她的朋友就是你的朋友；等等。

在给出理由来解释及为自身辩护时，你做了好几件事。你影响了人们解读你的心智、判断你的行为及谈论你的方式。你通过含蓄地承认所援引理由的规范效力来表达个人意见：促使他人期待你未来行为也受到相似理由的指导（如果并非如此则向你问责）。你也表明，在评价他人行为时，你用的理由与你援引来为自己辩护的理由是类似的。最终，在

你参与的一段对话中，他人可能会认可你的正当理由、提出疑问或援引出自身的理由，该对话有助于你与他人协作，而且从该对话中，实际上也会逐渐产生要共同遵守的规范。

把社会协作机制简化为遵守规范、“读心”或这两个机制的结合，便不能得知有多少人类互动旨在为自身辩护、多少旨在评估他人的理由（他人提供的理由或我们归于他人的理由）、又有多少批评过往或当前互动并预测未来互动。在这些与互动相关的互动中，理由是主要的。注

事实上，用来辩护的理由缩小了规范和“读心”之间的差距。当我们为自己辩护时，我们呈现出的动机符合规范，我们展现的规范也带有动机。换句话说，我们心理学化规范并“规范”心理状态。如果这样做，我们的目标就不是对自身行为及互动给出客观的社会学或心理学解释，而是通过维护并提升自身声誉、影响他人声誉来实现有利的协作。

通常在哲学、心理学和社会科学中，理由的社会协调作用能得到凸显。然而，人们普遍认为，理由归因是“读心”最复杂的形式。我们认为，就“读心”而言，理由归因通常具有误导性。它赋予理由的因果作用主要是虚构的；人们在选择归于自身或他人的理由时，与其说是基于这些理由的心理准确性，不如说是基于其作为正当理由的价值高低。我们认为，理由的解释性运用便是在实践理性的辩护功能，它把理由归结于个人，这样，充分的理由便不仅被视作对某个思想或行动的辩护，还是对这一思想的思想者、这一行动的施动者的辩护。

拿出及评估理由的能力进化不是为了提高心理学洞见，而是为了作为捍卫或批评想法及行为、表达承诺及创造相互期待的工具。理由归因的主要功能是为自身辩护及评估他人的正当理由。

沟通面临的挑战

不只人类合作，人类沟通也因其规模、多样性及机制的复杂性而分门别类。动物会用一小串不同的信号来传达每时每刻的基本信息：例如，出现天敌时发出警告声，出现竞争对手时发出恐吓声，出现交配对象时发出求偶声。人类语言拥有庞大的词汇库及强有力的句法，所以能制造各种语言表达。在语言的帮助下，人类能沟通关于任何可能话题的简单或复杂看法，如时空上相距甚远的事件、一般事实或抽象概念，既包括动物沟通中不可能出现的话题，还包括人文知识的热点话题。

人类迄今已拥有最丰富的代码，更多的是用这些代码进行沟通，而不是用来编码。正如我们注意到的，人们在说话时不能将其意思完全编码。他们所做的是为意义提供部分编码的证据。受众从这一语言证据及上下文信息来推断说话者欲表达的意义。例如，当阿兹拉（Azra）对马可（Marco）说“这个案例很棘手”时，马可以为她的意思是，他用来写侵权法论文的这个法律案例太棘手了，应该选个简单点的案例入手。如果这么解读，马可就远不仅是解码阿兹拉的这句话——其实在其他语境中这句话可以用来传达完全不同的意思——虽然如此，但马可的解读并没有超出我们一直使用的普通用语理解程序。

然而，保证话语被理解只实现了一半的沟通目标。说话者通常不仅想要所言被理解，还想要被信任（或被服从），换句话说，对其受众产生一定影响。受众通常不仅想要理解说话者的意思，还想通过这样做，学到某些处世知识。这需要受众不仅理解别人所言，还要认可它。阿兹拉告诉马可“这个案例很棘手”的意图是，让马可认可其暗含的建议。马可倾听她所言是为了从中得到某些指导。但是，马可也可能理解了阿兹拉的话但不同意。在人类沟通中，理解不代表认同。（例如，当你读某本书时，你会盲目地相信自己理解的任何事物吗？当然不会！）

传达的信息得不到认同，沟通对沟通者或受众都不会有好处。如果对二者都没有好处，沟通就不可能进化。那么，若受众盲目地认可传达的信息，能保证最佳的沟通效果吗？当然不能。想象一下，若人们盲目

地相信或去践行别人所言，沟通会变成什么样子呢？沟通者可能会很满意，因为他们有无限的可能去影响轻信又听话的受众。另外，不尽如人意的是，若受众认可别人说的任何话，就会深受各种错误信息和阴谋之害。

只有当沟通者值得信任时，盲目信任才会有意义。18世纪的苏格兰哲学家托马斯·里德（Thomas Reid）断言，上帝让人类具备这样一对性情：

> 这个聪明又仁慈的大自然创始人，认为我们应该是社会生物且应该通过他人提供的信息，接受自身知识中最伟大且最重要的部分，他为把这些意图根植在我们的天性中，创造了两个彼此吻合的性情。第一个性情是让人们讲真话……（第二个性情）是让人们相信他人是诚实的并相信他人所言。(注)

这个无所不能的造物主可能已经决定，要让人类既轻信又诚实。自然选择不大可能制造并维持这种完美和谐。如有可能，当欺骗轻信的受众符合沟通者的利益时，自然选择会鼓励沟通者毫不犹豫地这样做。同样，它也会鼓励受众不轻易上当受骗。

当信息发出者与接收者的生物利益即便不完全相同，但至少充分一致时，沟通仍然能在自然选择下进化。譬如蜜蜂的例子（刻板的工人通过帮助蜂后生产，促成蜜蜂基因的繁殖）。当只涉及共同利益的话题时，沟通才能进化。基于这些共同利益，沟通者欺骗受众不会有好处，受众怀疑传达的信息也不会有好处。雌狒狒让臀部形成巨大肿胀，是为了告诉潜在的交配对象，它们已经准备好交配。

在沟通关于共同利益的话题时，对话双方表现出诚实与信任都有好处。但人类沟通当然不只涉及共同利益的话题。人类可以任意发出语言信号来传递信息或误导他人。不像蜜蜂的摇摆舞或雌狒狒的臀部肿胀，

这些语言信号本质上并不可靠。人类不仅与至亲或合作伙伴沟通，还与竞争对手及陌生人沟通。每个人都会撒谎、欺骗。即便是孩子，在大约4岁的时候，就开始期待并践行某种程度的欺骗。[注]

人类坚持从与他人的沟通中获取大量信息。不进行大量沟通，他们甚至不能成为合格的成年人或过上正常人的生活。由于对沟通的依赖，人类既可能被欺骗，也可能被算计。尽管存在这种风险，沟通却进化成了人类生活中主要且必不可少的部分，这是为什么呢？

沟通是特殊的合作形式。正如我们在第九章中所见，合作进化通常会产生显著的问题。这种期待似乎是合理的，既能解释人类合作进化的理论，也能解释沟通进化。[注]这些理论真的能做到吗？

实际上，沟通是非常特别的情况。在大多数标准合作形式中，若骗子能侥幸逃脱，那么他们骗人就有好处。对于任何给定的个体而言，在家务中偷懒、工作中游手好闲或骗税等，如果没被发现，就会降低合作成本而不会折损收益。要是这样的话，合作只能在特定条件下进化并持续进行——特别是，当有序的监视及制裁让欺骗付出的代价通常高于获得的收益时，或当社会信息流会让骗子丢掉声誉及未来合作机遇时。然而，沟通进化必须满足的条件与这些特定条件大不相同。

尽管很难测量及测试，我们仍有理由相信，沟通者大多数时候都会自发地选择诚实，即便他们能轻易逃脱不诚实带来的后果，也会自发地选择诚实。这是为什么呢？因为诚实通常对于实现人们的沟通目标有用且有必要。人们通过沟通来与他人协作、组织协同行动并从乐于助人的朋友那里寻求帮助。在这种情况下，即便欺骗了对方没有被察觉，也不会带来任何好处；这只会是愚蠢的行径。

但是，在其他情况下，这些情形也相当普遍，一定程度的欺骗可能会带来好处。人们通常试图为自己塑造一个良好的形象。因此，即便那些自认为平时非常诚实的人，也会夸大自身的优点及成就，用虚伪的借

口来为其失败做掩护，做出不一定会去遵守的承诺，等等。有些夫妻，即便绝大多数时候都同心协力、真诚相待，但也会掩盖自身的出轨行为。在某些情况下，坦诚的沟通通常被视为无力之举，诚实反而会让人丢失许多机会，这些情况包括政治、广告、宣传及某些商业行为。虽然人们被人骗的频次可能比得知真相的频次要低，但是只要他们被骗，可能就要付出惨痛的代价。

一方面，我们相信，即便最真诚的人也不是永远都值得信任的。另一方面，有些人虽毫不担心撒谎要付出的代价，但最终仍选择讲出事实，而不是像在大多数场合那样撒谎，这并不是因为他们诚实，而是因为利益使然。处置骗子的标准方式是加以管束或将其踢出局，但这种方式可能不太适用于不得已撒谎的骗子。特定版本的以牙还牙策略[注]——你对我撒谎，我对你撒谎；或者你对我撒谎，我不听你的——造成的伤害可能会和惩罚撒谎者造成的伤害一样大，甚至可能更大。特定版本的“选择伙伴”策略[注]——你对我撒谎，我不听你的，我听别人的——如果进行系统性应用的话，会让我们错失许多有用的信息。

更为普遍的是，我们要尽可能从沟通中获利，同时最小化被骗风险，这需要用复杂的方式来过滤沟通信息。例如，当我们不假思索就怀疑曾在某一话题上撒过谎的人时，便没有考虑到，尽管这些人撒过谎，但他们是无人能及的百事通，他们在其他话题上通常还是可信的。而当我们不假思索就相信那些曾经可信的人时，便没有考虑到，在某一新的话题上，当欺骗之举对他们有利时，我们很可能就会被骗。因此，我们对沟通者传达的信息不能一概而论，而要视情况来给予信任，这就要求我们不仅要仔细考虑对沟通者过往表现的记录，还要考虑沟通的情境和内容。

认知警觉

虽然沟通对人类非常有利，却让其很容易受到错误信息的攻击。因此，我们在之前的研究中就已表明，[注]必须要施加强有力且不间断的压力来发展一套“认知警觉”机制，使之能够适用不断调整的信任。虽然没有故障保险方案来精确校正信任，却可以利用许多相关因素来校正信任。对我们而言，重要的是获得沟通的好处而不成为其受害者。因此，任何机制，只要能高效地促进我们的信任分配，就很可能已被我们利用。

认知警觉所涉机制分为两类：一类机制把注意力集中于信息源并帮助回答该相信谁的问题，另一类机制则把注意力集中于内容并帮助回答该相信什么的问题。

我们应该在什么时候相信谁，关于什么话题及问题？我们从医生或水管工那里得到不同的建议。我们更容易相信的建议出自不涉及个人利益的目击者，会在人们迟疑不决或过于坚持时提高警惕，会考虑到人们诚实及能力方面的风评，等等。此外，最近许多实证研究表明，儿童很早就开始显露考量沟通者能力及慈爱程度证据的能力，据此来决定该相信谁。[注]

对沟通者本人的信任并不能代表一切，因为并非他们所有的话都值得相信。一方面，没人能说服你相信2+2=10、说服你相信月亮是一片奶酪，或说服你相信自己还未出生。若这些话是世上你最相信的那个人告诉你的，你不会按字面意思来理解它们，而是会等待解释（例如，四进制下，2+2=10；月亮的确像一片奶酪；从宗教意义上来讲，你的确还未“再生”）。另一方面，若你在世上最不相信的人对你说“正方形的边比三角形的边多”“巴黎是法国的首都”“你活到这个年纪一定吃了许多食物”，尽管你不相信这个人，仍会同意其观点。

不论谬论出自谁口，人们都不可能直接相信其字面意思。同样，不论真理出自谁口，人们都不可能怀疑其真实性。不过，大多数信息都处

于这两个极端之间：正确性不明显，错误性也不明显。当我们判断大多数断言的可信度时，都是基于这些断言与我们自身信念的契合度。若我们发现某一断言与自身信念不一致，很可能会拒绝接受该断言。但是，若拒绝接受质疑自身信念的断言，可能会错失适当修改自身起初便有误或现在需要更新的信念的机会。

对沟通者本人的警觉和对其话语内容的警觉可能指向不同的方向。若我们相信沟通者本人，而其断言却与我们自身信念矛盾，这时，出现信念修改是难免的：若我们接受其断言，则必须修改自身与之矛盾的信念；若我们拒绝接受其断言，则必须修改认为此人可信的信念。

尽管认知警觉还做不到故障保护，却能让我们作为受众，根据所传达信息的可靠性来进行梳理，通常还能让我们从这些信息中获利。另外，对沟通者而言，受众的警觉使他们的欺骗收益减少——受众的警觉降低了他们成功的概率——而且如果他们被抓获，便增加了成本，会让他们的信誉及声誉受损。难道我们不应该不仅对他人传达的信息有所警觉，还要对自身信念形成机制的产物有所警觉吗？当我们谈及源于自身认知过程的信念时，我们自己便是沟通者。我们的认知机制慢慢进化，来为我们提供服务。不像他人，我们的这些认知机制真的没有自身利益。此外，我们还认识到，已经有一些元认知策略，可以用来调节我们对自身记忆、知觉及推论的自信。因此，当我们自己就是沟通者时，还认为警惕沟通者更有利，便几乎没有理性可言。

对沟通者的警觉就是这样吧，那么对沟通内容的警觉又如何呢？为更好地决定应该认可或否决哪些新信念，或更好地决定应该修改哪些旧信念，检查新认知及推论与旧信念之间的一致性难道没有用吗？实际上，一致性检查既不是简易的步骤，也不是误差检验。而且，若一致性检查值得做的话，难道其产物本身不应该接受两重、三重甚至更多检查吗？若你乐意怀疑自己，难道不应该怀疑自身的怀疑吗？我们知道，没有证据表明，人们会警惕自己的认知及推论内容（除了在特殊情况下，

当他们有确切的根据认为其官能或直觉没能正常运转时）。我们也知道，没有充分的论据表明这样做真的有利。更为普遍的是，人们似乎不在乎自身话语的前后不一致，除非某些事（通常是某些人）让他们不得不去在乎。

理性的论证功能

警觉是有代价的。即便警觉总的来说是有利的，也还是会让人错失机会。譬如，既然你不确信树林里的这些蘑菇是否可以食用，不去采摘便是明智之举。即便后来得知它们是美味的鸡油菌，当时选择不去采摘，也同样不失为明智之举。认知警觉作为警觉的一种形式，同样也要付出代价。因为我们没有给予传信者充分的信任，便很可能否决有价值的信息，这种选择显然是错失机会。

不论认知警觉多么有用或多么必要，都阻碍了信息流动。但是，对信息的受众而言，精心校准的认知警觉带来的好处与付出的成本相比要多得多。另外，对沟通者而言，受众的警觉似乎只会让他们付出较高的代价。这种警觉不仅阻碍了不诚实的沟通者，也阻碍了诚实的沟通者。诚实的沟通者可能热切地想去传达真实且中肯的信息，但在其受众看来，他还没有充分的权威来让他们接受他传达的信息，在这种情况下，二者都丢失了机会——沟通者丢失了影响受众的机会，受众则丢失了获得中肯知识的机会。

你从给定来源中获取的信息越显得中肯，你越应该警觉，但这一警觉造成的代价也越高。下面是一个戏剧性的历史案例。理查德·佐尔格（Richard Sorge）是苏联间谍，第二次世界大战期间在日本做卧底并为苏联人提供了许多有价值的情报。然而，当他告诉苏联人纳粹德国将于1941年6月侵略苏联时，斯大林竟否决了其报告："就是这个杂种……草率报告了德国进攻的日期为6月22日。你认为我应该相信他吗？"[注]事后

回想，这是多么愚蠢的错误呀！但是，若这一情报正确的话，则会非常珍贵；若有误却被当作正确的话，结果也会十分惨烈。正因为如此，有必要格外警觉甚至是怀疑。在我们的个人生活中，也会对别人的警告或许诺持怀疑态度，正是因为这些警告和许诺太中肯了，反而让我们不能不加深究地予以认可。

在此，推理便要发挥作用了。理由的论证性运用帮助真实信息越过认知警觉对社会信息流动设置的障碍。对受众有利的是，允许其更好地评估可能有价值的信息，虽然他们不会不加深究地认可这些信息。对沟通者有利的是，允许其说服谨慎的受众。

要理解论证如何帮助克服信任的局限，应考虑到，当沟通者不加论证就做出断言时，会发生什么。譬如，拥有权力的人通常希望人们认可其所言，仅仅是因为出自他们之口。这一期待通常是错误的：他们可能具备强迫他人服从的权力，却没有强迫他人深信不疑的权力。在某些文化环境中，陈述没有任何论据支撑的观点也比较常见，论证及通常的“聪明过头”不被提倡。在这种情况下，就智力这个术语的所有意义来看，不会蓬勃发展。在这两种情况下，信息流动都受到了阻碍。当允许论证时，这种流动会容易些。

作为沟通者，我们与受众沟通，若这些受众不会不加深究地相信我们所言，则会去审查我们所言与其自身信念的契合度。既然我们有时也会充当受众，便能够理解在沟通中受众如何评估我们所言。我们从这一理解中获利，并相应地调整所传达的信息。除非我们试图欺骗受众，否则便不必把其警觉视作对我们交际目标的阻碍。相反，我们也许能够在某种程度上利用受众的警觉，使之对他们和对我们自己而言同样有利。

一开始，即便受众不愿认可我们的主要论点，仍会不加深究地认可我们提供的相关背景信息。通过提供这类信息，我们可以拓展共识，基于此共识，人们会评估我们让人有点费解的断言。例如，

（门铃响了）

恩里科（Enrico）：一定是艾丽西娅（Alicia）或沙利文（Sylvain）。

米歇尔（Michelle）：沙利文其实不在城里。我肯定是艾丽西娅。

恩里科：是的！

恩里科可以不加深究地相信米歇尔在此提供的信息，因为如果米歇尔不知道“沙利文其实不在城里”这个事实，是不会这样说的。然而，这一信息便是促使恩里科修改其猜想的理由。

说服受众的较好方式便是积极帮助他们检查你的断言与其信念（包括他们刚完全认可的你的观点）的一致性，或者如有可能，更好的途径便是帮助他们意识到，鉴于其自身信念，拒绝接受你的断言相比认可你的断言并不占优势，反而会让他们变得更不一致。换句话说，作为沟通者，你与警觉的受众沟通，当你诚实地展现话语的一致性时，你的可信度便可能会增加，因为受众无论怎样都会审查这些一致性的。合理的论据恰恰在于展现受众能独自进行评估的一致性关系。

作为受众，当你不仅听到某个断言，还听到支撑该断言的有利论据时，你可能会（直觉地或反思地）评估该论据，而且如果你认为该论据合理，可能最终会认可该论据及断言。沟通者的论据对你的好处表现在两个方面：一是通过表现这种你可能必须独自进行评估的一致性，使你评估该断言变得容易得多；二是如果这种评估最终促使你认可了相关信息，沟通会变得更有利。

沟通者要如何表现一致性呢？他从其受众已持有的信念中搜寻（或将不加深究地认可）支撑其断言的理由。在简单的情形中，此举可能需要用到仅一个论证步骤，就像下面本（Ben）和克瑞莎（Krisha）

（唐、茱莉娅和玛丽的两个邻居）的对话表现的那样：

克瑞莎：唐刚刚打电话了。邀请我们与他、茱莉娅和玛丽在他家共进晚餐。

本：玛丽也在那儿？那么她一定已完成了那篇必须要写的论文。

克瑞莎：如果她真的完成了，我会很惊讶。

本：当然，如果她没有完成论文，就会在图书馆写到很晚。

克瑞莎：但是你早上告诉我图书馆今天会很早闭馆。

本：啊，是的，我忘了。所以你说对了——玛丽可能还没有完成论文，毕竟她一直待在家里。

克瑞莎在此利用本提供的信息（图书馆今天会很早闭馆）作为理由来怀疑本“玛丽一定已完成了论文”的断言。克瑞莎的论据告诉本，他的观点前后不一致，他至少应修改其中一个观点。

要让受众意识到，若与反对沟通者的断言相比，认同沟通者的断言对其而言更具一致性，可能需要好几个论证步骤，正如下面这个例证展现的那样，这一例证基于我们在第一章中所遇的鸽巢问题（这个问题仅此一次涉及了真正的演绎推理）。

迈拉（Myra）：鲍里斯（Boris），既然你喜欢猜谜，那我说一个让你猜。在丹顿的村庄里住着22个农民。每个农民至少养了一头奶牛。任何农民养的奶牛数都不超过17头。那么至少有两个农民恰好养着同样数量奶牛的可能性有多大？鲍里斯，对此你怎么看？

鲍里斯：好吧，我认为有可能。

迈拉：我认为是肯定的！

迈拉和鲍里斯得出了不同的结论，但他俩都没有权威性，不能仅通过说一句“就是这样”来说服另一个人。虽然没有可以有力论证观点的权威，但可以求助于论据。

迈拉：好，想象一下，你来到村庄里，把这22个农民聚到一起，然后要求他们以小组为单位来站队，且每个小组里的农民都恰好养相同数量的奶牛。

鲍里斯：那有什么用呢？结果可能是没有谁恰好有相同数量的奶牛，那么，在你的每个“组”里，只会有一个农民。

其实鲍里斯的回复包含他还没意识到的蕴含条件，但迈拉强调了该蕴含条件。

迈拉：这不可能。既然任何农民养的奶牛数都不超过17头，那么小组数便不可能会超过17。养1头奶牛的农民组成一个组，养2头奶牛的农民组成一个组，以此类推，直到养17头奶牛的农民组成一个组。但到那时，鉴于有22个农民，且至多分成17组，那么至少有一组由好几个农民组成，不是吗？

鲍里斯：是的，然后呢？

迈拉：既然同组中的农民会养相同数量的奶牛，那么一定至少有两个农民恰好养相同数量的奶牛。正如我说的，这是确定无疑的，而不只是可能。

鲍里斯：你说对了。我现在意识到了。

迈拉所做的是向鲍里斯陈述一些直觉论据，这些直觉论据详细说明谜题所描述情境的某些明显暗示。不过，最后一个直觉论据所支撑的结论直接否定了鲍里斯起初的结论。当鲍里斯意识到这个问题时，对他而言更具一致性的是认同迈拉的看法并改变自身看法。迈拉和鲍里斯的对

话是简化的苏格拉底法则（Socratic method）：帮助对话者在其自身信念中找到理由来改变其看法。当推理反思首次出现在西方传统，即柏拉图（Plato）和亚里士多德的作品中时，处于核心地位的是理由的论证性运用，是在试图说服对话者、全体审判人员或立法机构时。社会层面的推理不言而喻。我们可以把苏格拉底式推理视作最优秀的推理。不过，尽管推理研究在亚里士多德的作品中早已出现，却朝着不同的方向发展。

“什么使论证有效”这一重要的规范性问题，促成了更为抽象的推理方式，并促进了纯逻辑的发展。现在只能把理由在论证中的运用视作更具一般性的推理原则诸多应用中的一个实际应用。经典推理者的新形象便显现出来了。与苏格拉底试图说服其对话者和对话者感受其论据的说服力不同的是，推理者的典范变成了科学家（很少是女性科学家）独自推理以更好地了解世界。从推理心理学的角度来看，这一论证是不合理的：弄混了推理（包括科学推理）在哪种程度上是社会活动、在哪种程度上是基于直觉。㊟

理由不是随意的修辞工具。如果是的话，要如何具备说服力呢？理由受到那些基于真正认知能力的直觉的支撑。直觉上合理的理由更可能支撑真实结论。譬如，想象一下，如果迈拉没有尽力说服鲍里斯，那么鲍里斯便会尽力说服迈拉认可其回答的正确性。鲍里斯成功的概率有多大呢？不是很大。其理由直觉上就不够充分。实际上，在尽力准备理由时，他有可能已意识到自己的理由不够充分。在这个案例中，令迈拉和鲍里斯产生分歧的是一个有明确答案的逻辑问题。即便令两人持有不同看法的这个问题没有明显正确的答案，但相比之下，其中某个答案可能会有更充分的直觉论据来支撑。当然，在某些情况下，两个自相矛盾的结论都能受到貌似可信但非决定性理由的支撑。在这种情况下，通常不能单凭论证改变人们的看法。

当我们试图说服他人时会拿出论据，或前瞻性地说，当我们认为也

许要这样做时会拿出论据。我们评估他人提供的论据，将之作为工具（虽有缺陷但仍特别有用）来认可合理的想法、摒弃不合理的想法。我们有时作为沟通者，有时作为受众，既能拿出呈递给他人的论据并从中获利，也能从评估他人呈递给我们的论据中获利。推理需要两种能力，一种是制造论据的能力，另一种是评估论据的能力。这两种能力相互协调且必然一起进化。我们认为，它们一起构成了理性的两个主要功能之一及推理的主要功能：论证功能。注

主要功能、次要功能和海龟

本书的中心论点是人类理性有两个主要功能，与人类互动的两个主要挑战相对应：理由归因主要充当辩护功能，推理主要充当论证功能。

为什么我们要说两个主要功能，而不是说只有这两个功能呢？坦率地讲，后者不够严谨。虽然我们不确定理性模块除了这两个功能之外是否还有其他功能，但仍想要给可能性留有余地。

譬如，我们不否认个人理由归因可能出于解释而非辩护的目的。有人在进行理由归因时，可能并没有评估这些理由的辩护价值。毕竟，这就是历史学家在力图用无偏见的方式解释从古至今的人类行为时所做的。难道理由的这种解释性运用没有好处吗？难道它不能至少作为理由归因的次要功能吗？

不论是拿出说服他人的论据，还是评估他人提供的论据，我们不否认推理能单独进行。难道这种单独的探究式推理没有好处吗？即便推理的主要功能没有产生这些好处，难道不应该至少是推理的次要功能产生了这些好处吗？

好吧，也许吧，但为了更好地了解形成这种断言需要些什么，让我

们看一下海龟的例子。

虽然海龟是陆地爬行动物的后代，但几乎每时每刻都生活在水中。雌海龟要上岸产卵，在沙滩上挖出巢穴，这些巢穴由此便成为所有海龟的出生地（之后这些小海龟会尽可能快地爬回水中）。

海龟的四肢已近乎完美地适应了海里的生活：其前肢进化成了鳍状肢，后肢进化成了尾鳍。怎样能更好地旅游呢？海龟运动的一般模式是怎样的呢？不过，不止一次，雌海龟必须利用同样的肢体爬出水面，在沙地上挖巢穴。虽然这种运用四肢的方式罕见且笨拙，却对成功繁殖有直接作用，而且明显是适应性的。然而，为了最终证明挖巢穴是海龟四肢的次要功能，我们必须找到其四肢的特点，而挖巢穴已成为其四肢特点的最佳解释。否则，便最好把海龟四肢挖巢穴的功用描述为意外的好处。注

雌海龟挖巢穴的本领很显然是有好处的，尽管其使用的四肢不能专门适应这项任务，尽管其表现格外笨拙。同样，如果理性有不同于其两个主要功能的有利影响，那么这些有利影响可能相当于次要功能。要是这样，我们希望找到理性用以实现这些效益的特点。不然的话，我们可能会怀疑这些仅是理性的负面影响。

貌似可信的是，通过理由归因来解释人们想法及行为的能力基本上是有利的（即便正如我们看到的，它充其量只是对所涉心理过程的一种曲解）。真正的个人探究式推理（与精神模拟论证截然相反）在指导信念及决策方面利大于弊，这一点并不那么显而易见，但也不是匪夷所思的。然而，即便假定理性的这两个用处通常都是有利的，严格来说，这也不意味着它们是理性的功能。要证明它们是理性的功能，一方面，我们必须找到理由归因的某些具体特点；另一方面，必须找到推理用来产生这些特定效益的某些具体特点。我们没有找到任何此类证据。此外，推理不仅缺乏用来实现其所谓的探究式功能的专门化特点，其好几个显

而易见的特点实际上还会削弱这一探究式功能。

关于理性可能具备的次要功能，我们当然不应该抱有成见，但目前的挑战是，要确定到底哪些有利影响可以解释理性进化的原因——换句话说，确定理性的主要功能。在试图应对这一挑战时，我们否认传统理性这类有能力在各个领域获得更丰富知识、做出更明智决定的事物曾经进化过。我们认为，进化了的是更适度的理性模块，即许多模块中的一个直觉推论模块，该模块在其两个功能（辩护及论证功能）的帮助下专门用来产生对理由的直觉。

在本书第四部分和第五部分，我们会用丰富多彩的证据证明，理性恰好适用这两个功能。此外，我们还会展示，这种新颖的互动理论如何阐明理性在人类事务中发挥的作用。

---

1. Darwin 1981, p.46.
2. Darwin 1981，第188—189页。
3. 可参见：Einborn & Hogarth 1981; Nisbett et al. 1983。
4. Cosmides & Tooby 1987, 1989; Tooby & Cosmides 1989.
5. Cosmides 1989.
6. Girotto et al. 2001; Sperber, Cara & Girotto 1995.
7. 欲了解所用方式不同却表达相同意思的论据，可参见：Kruglanski & Gigeren-zer 2011。
8. Heyes 2003; Sterelny 2003.
9. Mercier & Sperber 2011.
10. Bicchieri 2006.
11. Malle 2004.
12. Reid 2000, pp.193-194.
13. Mascaro & Morin 2014.
14. 欲了解最近有关合作与沟通关系的观念，可参见：Godfrey-Smith & Martínez 2013；Skyrms 1996, 2010；Sterelny 2012；Tomasello 2010。

15. Axelrod 1984; Blais 1987.

16. Baumard, André & Sperber 2013.

17. Sperber et al. 2010.

18. Clément 2010；Harris 2012；Mascaro & Sperber 2009.

19. Montefiore 2003, p.360.

20. 欲了解经典逻辑史中对其辩论根源的解读，可参见：Dutilh Novaes 2015。

21. 我们不是第一个形成论证推理理论的人。其他人（如Billig 1996; Perelman & Olbrechts-Tyteca 1958; Toulmin 1958）一直认为推理主要是关于提出及评估论据的。他们主要是基于内省的依据并从哲学的视角来实现此举的。我们的做法可能更新颖一点（Mercier 2016a; Mercier in press; Mercier & Sperber 2011），因为我们采用的是实证依据及自然主义的进化视角（也可参见：Dessalles 2007, 在其形成的不同的进化视角下，论证发挥着重要的作用）。

22. 欲了解适应特性与有益的副作用之间的差异，可参见：Sober 1993; 欲了解更多海龟四肢进化方面的信息，可参见：Wyneken, Godfrey & Bels 2007,pp.97-138。

第四部分

理性有所为、有所不为

在第七到十章中，我们详尽阐述了新颖的互动推理理论。根据这一理论，理性的功能是在我们与他人的对话中制造并评估正当理由及论据。而从标准的唯理智理论来看，理性的功能则是让某人独自形成更好的信念、做出更好的决定。我们会发现，这两种理论对理性功能的预测对比鲜明，而且我们现在能够对其预测进行检验。理性是客观的还是有偏见的？是严苛的还是怠惰的？会帮助孤独的推理者吗？在互动语境下，其产生的结果是更好还是更糟糕呢？对脑海中浮现的这些问题，下面举出的例证可以使人们进一步了解理性的前景及其危险。

第十一章
推理为何存有偏见

人们对检测适应性假设的方式有许多误解。为了解理性的作用及进化的原因，我们似乎必须利用考古学或遗传学数据，来弄清楚我们祖先的推理方式。幸好，百万年前的颅骨及基因测序，对于检测有关进化机制功能的假设并非不可或缺。想一想：我们虽然不知道人类眼睛进化的时间及方式，却知道其用处。重要的是，器官或认知机制的功能与其构造及影响基本一致。眼睛的特点对其功能有用吗？基本上是。眼睛本身能很好地发挥功能吗？基本上能。

我们可以用同样的逻辑来指导对人类理性的数据检测。人类理性的特点能完美地践行唯理智理论或互动理论假定的功能吗？哪一个功能践行得更好呢？在大多数情况下，理性不能同时践行好这两个功能，所以下面会有大量例证来帮助我们选择。我们会先运用一个历史案例来讲述理性到底能还是不能践行好这两个功能。认证的科学天才是如何推理的呢？理性能帮助他摒弃错误的信念、得出更合理的结论吗？

莱纳斯·鲍林（Linus Pauling）30岁时便荣获美国化学学会理论化学奖，那时，一位资历较深的同事就预言他将获得诺贝尔奖。[注]结果，鲍林居然获得了两个诺贝尔奖——诺贝尔化学奖及和平奖，成为继玛丽·居里（Marie Curie）后又一位获此殊荣的科学家。他埋头研究DNA结构，是詹姆斯·沃森（James Watson）强有力的竞争对手，但还是与第三个诺贝尔奖失之交臂。的确，詹姆斯·沃森一直很担心鲍林再次萌生“惊人的想法”[注]。鲍林从前的学生把他描述为“上帝般的、超人似的伟大人

物”[注]，其学生后来也荣获了诺贝尔奖。毋庸置疑，莱纳斯·鲍林想法超群，而且对最严格的科学惯例的研究造诣颇深。

在获得诺贝尔奖之余，鲍林有时会忖度维生素C的功效。起初，他主张每天注入强剂量的维生素C，称这样可以预防感冒和其他一些小疼痛，但不主张把维生素C作为治疗严重疾病的药物。他的这一主张在遇到伊万·卡梅伦（Ewan Cameron）后发生了改变。伊万·卡梅伦是一个外科医生，曾做过一项小型研究，研究发现维生素C对治疗癌症有积极作用——至少看似如此。鲍林和卡梅伦合力证明维生素C治疗癌症的潜力。多亏鲍林的影响力及其坚持不懈的努力，享有盛名的梅约诊所（Mayo Clinic）才同意开展一次大规模的、严格把关的试验。[注]不幸的是，对鲍林、卡梅伦及癌症患者而言，结果是负面的。维生素C没有任何治疗癌症的作用。[注]

这时，鲍林本可以客观评估这些例证：边缘理论及没有严格控制的小研究，医疗共识及严格控制的大诊所试验。基于这些例证，绝大多数研究人员及医生都得出结论，维生素C治疗癌症的作用没有得到证实。鲍林却没有进行客观推理。他杜撰了一种特殊情形。鲍林认为，梅约诊所的第一项研究失败是因为没有挑选合适的参与者。当针对该问题开展第二项研究时，[注]鲍林又指出另一个问题：新患者没有“无限期地”[注]服用维生素C。任何标准的癌症研究都不会提出鲍林这样的要求，鲍林的要求只有卡梅伦开展的这个小试验才能符合。

鲍林最过分、最有偏见的推理几年后出现在其发表的一篇文章中，他在该文中提出了癌症治疗临床试验的三个评价标准。[注]虽然“对许多癌症病患进行临床试验所获得的大多数试验结果都符合这些有效性标准”，[注]但仍出现了一个例外情况，“有个临床试验结果不符合任何一个有效性标准”。[注]你能推测一下这个异常试验是什么吗？这项研究“被描述为维生素C（每天10克）与乳糖安慰剂的随机双盲对比”研究（着重强

调）。[注]在鲍林看来，上百个样本中唯一存在错误的癌症研究便是梅约诊所的研究，因为这项研究让鲍林在整个科学界难堪。

我们没有理由控诉鲍林是有意欺骗。因为他本人每天都在服用高剂量的维生素C，其妻子也使用同样的养生法来抵御（虽然没有成功）胃癌。但是，大多数人看到的是，鲍林对维生素C疗效的评估是基于精心挑选的证据及有偏见的论据。即使鲍林多年来坚持服用高剂量的维生素C，但还是被确诊患了癌。他不承认自己失败了，声称要是没有维生素C的支撑，他早就患癌了。[注]

大多数受人景仰的科学家都犯过错，但鲍林的错误愈加严重，他持着非正统的信念越走越偏。但是，他的推理方式并不罕见——任何对科学家有所了解的人都可以做证，科学家并不是践行客观性的典范（在第十八章中会详细论述）。毋庸置疑，即使是最伟大的科学家，也会用有偏见的方式进行推理。

偏见总是不好的吗

我们要如何运用认知机制来形成合乎情理的信念呢？让认知机制不带偏见仿佛是一个可行的方案。偏见之所以会招致许多批评，部分原因在于其通常被定义为“赞成或反对某个人或某个团体的倾向或偏见，尤其以一种不公正的方式”。[注]然而，心理学家对此却持有不同的态度，他们通常用这个术语来指犯某特定类型错误的系统性倾向。这类错误不必影射任何道德、社会或政治问题。

存在第一类偏见只是因为认知有加工成本。人们可以使用启发法来降低成本，这种认知捷径通常比较可靠，但有时还是会失误。特沃斯基和卡尼曼[注]合作研究的“可得性启发法”（availability heuristic）为此提供了一个范例。该方法利用事件浮现在脑海中的容易程度来猜测其实际

出现的频次。譬如，当研究员在实验中问参与者“字母R作为英语单词的第几个字母出现频次更高，第一个还是第三个？”时，大多数参与者都回答其作为第一个字母出现的频次更高，而事实是作为第三个字母出现的频次更高。参与者利用启发法试图回忆以R开头的单词［如“河流”（River）］和将R作为第三个字母的单词［如“无聊的”（Bored）］，并假定两类单词在脑海中浮现的容易程度便反映了各自实际出现的频次。字母R（及英语字母表中其他7个辅音）的情况比较特殊，可得性启发法碰巧行不通。而对于英语字母表中的其他13个辅音字母，运用同样的启发法却能得出正确的答案。

尽管可得性启发法存有偏见，且的确会衍生系统性错误，但当人们考虑用另一种方法，即在自己的知识储备中试图搜寻出所有将R作为第一个字母或第三个字母的单词时，可得性启发法的实用性还是显而易见的。虽然启发法有时并不好用，但我们对这点并不担心，我们更担心参与者运用这一艰苦的思索过程，只是为了回答某个心理学测验。此外，心理学家歌德·吉仁泽（Gerd Gigerenzer）及其同事指出，我们在许多情况下运用这种启发法不仅不怎么费力，而且较之更复杂的策略，运用这种启发法反而能得出更好的结果。[注]吉仁泽认为，启发法不是“权宜之计”，而是“高效”且非常可靠的思维方式。

鉴于错误有大有小，第二类偏见便出现了。[注]我们应该更谨慎些，不要犯危害太大的错误，平常犯一些无伤大雅的错误还是可以原谅的。

下面举出一个简单明了的例子来阐明犯错代价的失衡如何造成了适应性偏见。熊蜂的认知机制帮助其躲避天敌。其天敌之一便是蟹蛛，这种小蜘蛛会在熊蜂采蜜时捕捉它们。某些蟹蛛会停歇在花朵上，并伪装成花朵的颜色，这样便隐藏起来了。为更多地了解熊蜂躲避隐藏天敌的方式，托马斯·英格斯（Thomas Ings）和拉斯·奇卡（Lars Chittka）合作发明了小型机器蜘蛛来做实验。[注]他们让所有的机器蜘蛛都停歇在黄色花朵上，把某些机器蜘蛛设置成白色的（没有隐藏的），而把另一些设

置成黄色的（隐藏的）。为模拟觅食风险，英格斯和奇卡还造出了小型螯，当熊蜂停歇在有“蜘蛛”的花朵上时，他们便用这些螯将其在两秒内俘获。

在实验的第一阶段，他们把熊蜂分成了两个组，让一组面对隐藏的蟹蛛，另一组面对没有隐藏的蟹蛛，这两组熊蜂都有许多机会飞到花朵上，来探察其面对的是哪类天敌。出人意料的是，即使蜘蛛隐藏起来了，两组熊蜂也都很快就学会了躲避藏有蜘蛛的花朵。保护色并非不起作用：有些熊蜂在面对伪装的蟹蛛时，要花近两倍的时间来检查每朵花，才能探察到隐藏的蟹蛛。这个例子表明，必须花费对身边环境形成精准表征的成本，即用来探察蟹蛛的时间，而在这期间不能采蜜。

熊蜂误停在藏有蟹蛛的花朵上的成本（高成本）与不必要地躲避无蟹蛛隐藏的花朵的成本（低成本），二者之间也存在不对称性。这一不对称性也影响了熊蜂的行为。在实验的第二天，这些熊蜂便认识到了周围环境存在的觅食风险。它们没有不断花时间（0.85秒）探察每朵花，来确保没有蟹蛛藏在上面，这样的话，那些停歇在藏有蟹蛛的花朵上的熊蜂，则面临较高比例的“错误情报”：与其他的熊蜂相比，这些熊蜂更可能去躲避实际并无蟹蛛隐藏的花朵。

这项实验阐明了一种十分微妙的方式，即便是相对简单的认知系统都会用这一方式来调整耗费在认知任务（探知隐藏在花朵上的蟹蛛）上的时间和精力，使之一方面适应任务难度，另一方面适应假阴性错误（明明藏有蟹蛛却假定没有）和假阳性错误（明明没藏蟹蛛却假定有）的相对成本。这种成本差异会导致偏见产生，让人犯更多的假阳性而不是假阴性错误。不过，这一偏见还是有好处的。

从惊奇到证伪

虽然熊蜂的例子表明，偏见可能有好处，却谈及了十分具体的成本，即觅食成本和错失觅食机会的成本。存在更普遍适用的认知偏见吗？

认知机制的主要目的是准确表征生物环境，至少是生物环境的相关层面。据称，只有在未来，尤其是在不远的未来，对生物环境的准确表征对生物才有直接相关的意义，它帮助生物预测潜在的危险，还帮助其改造周边环境来保护自己。对先前和当前环境的表征可能也是相互关联的，但只是一种间接关联，它们可以为未来环境的表征提供唯一可用的证据。也可以说，认知的主要目的是让生物准确预测接下来可能会发生的事。认知机制应格外关注任何与其当前期待相悖的信息，并利用这一信息适当修改当前期待。若信息与期待不同，至少会让人类产生典型反应：惊奇。这种惊奇的体验是认知资源的突然调动，来调整刚遭到挑战的期待。

的确，我们总是格外注意出乎意料之事，这一习惯根深蒂固。因此，当我们看到别人对出乎意料之事不怎么惊奇时，我们就该惊奇了。假设你朋友乔（Joe）遇到了两只长得相像的猫，其中一只猫在叫，另一只猫在说话，而他却没有特别关注那只说话的猫，这时，你会怀疑他有点不正常。即使一岁的婴幼儿都希望别人注意到他们的惊奇。当他们看到惊奇的事物时，便指向那个东西，来让一旁的大人也注意到。他们会一直指着那个东西，除非得到了合适的响应，或因大人不理不睬而不再指望大人能注意到。[注]

这并不意味着人们就会寻求意外的事。如果你足够幸运，身处环境良好，你便能预测未来走向。若预测出现偏差，就会发生意外的事，而这些意外的事很可能就是坏消息（因而有句伪造的中国古语“宁为太平犬，莫做乱离人”）。不过，即便你不喜欢意外的事，若环境已蓄势待发，还是尽早获悉此事为好。被解雇很糟糕；在真的被解雇之前就得到了相关信息，并先为之意外一番会更好些。因此，与提供重复信息的沟

通者相比，提供意外信息的沟通者会得到人们更多的青睐（假设他们的可信度是一样的）：人们会购买大标题为“人咬狗”的报纸，而不是大标题为“狗咬人”的报纸。

如果理性是个人认知的工具，旨在让人们形成更好的信念、做出更好的决定，那么理性应该格外关注与我们自身期待相悖的信息。理性应该为我们当前结论寻找反对论据，应该找出理由来质疑我们当前的决定，找寻与我们当前信念相冲突的记忆。与没能关注正面证据及论据相比，没能关注负面证据及论据通常要付出更高的代价。

认识到这一点非常重要，因为学者反复告诫说，这是形成合理信念的最佳方式。12世纪，罗伯特·格罗斯泰斯特（Robert Grosseteste）主张，要杜绝错误假设，就应该先提出一个假设，然后再看是否任何遵循该假设的事物都不合逻辑或与事实相反(注)。400年之后，弗朗西斯·培根（Francis Bacon）推翻了格罗斯泰斯特隶属的学术传统，把观察放在其哲学体系的核心位置。尽管如此，仍然有人认为要寻找反例，认为观察不应该是“幼稚地”枚举符合研究者假设的案例。学者要专门找寻能证明其假设有误的例子：人们必须“适当拒绝或排斥自身假设，以分析这些假设的性质”。(注)

又过了400年，这回是卡尔·波普尔（Karl Popper）来挑战当时的主流观点，即培根哲学体系的遥远分支。为此，波普尔比以往任何时候更强调证伪的重要性。他认为，科学之所以与其他知识区分开来，不是因为科学理论可以得到证实，而是因为科学理论可以被证伪：原则上，只要单次观察到（根据某个给定的理论来说）某个永远不可能发生的事件，就能证明该理论有误。当然，波普尔也认可，在实践中，单次观察还不足以促使科学家摒弃某个理论——毕竟，观察本身就可能有误。但是，若有众多观察证明某一理论有误且这些观察足够可信，那么要是不摒弃这一理论的话，至少应该对此予以修正。科学家提出某一理论时就想到了会被证伪。他们不断完善自己的理论，寻找会证伪自己理论的证

据，摒弃已被证伪的理论，坚持已经反复检验的理论。

尽管这些学者的观点存在差异，他们仍一致认为，反例及其他与期待相悖的信息让我们摒弃错误的信念，在知识的积累中发挥关键性作用。正如波普尔所言："在追寻真理的过程中，最好的计划莫过于一开始便批判自身最珍视的信念。"(注)

证实偏见

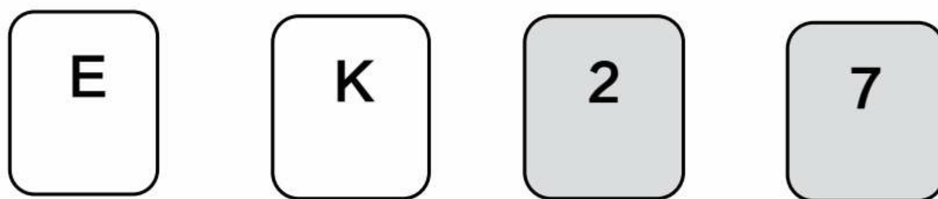
彼得·沃森在伦敦大学学院设计的实验任务就颇为机智，而且其实验地点与其执教的伦敦经济学院只有几步之遥（当然不只是地理位置接近），该实验任务预计会对推理心理学产生一定影响。沃森拓展了波普尔对科学理论的见解，将其延伸到日常推理中，提出问题：人们是靠证伪来形成更好的信念吗？

我们在第二章中已对选择任务进行了论述（不清楚可以看看图11.1），而该任务恰恰是为了模拟科学思维的方方面面。沃森认为，参与者选择的卡片必须能作为检验规则的适当证据，就如科学家必须要找到检验其假设的适当证据一样。如我们所见，在这个例子中，正确的做法是选择写有元音E和奇数7的卡片，因为只有这两张卡片可以证明以上规则有误（若元音E的另一面是奇数，或奇数7的另一面是元音便可证明）。然而，大多数参与者却没有选择写有奇数7的卡片。许多人都选择了写有偶数2的卡片，想必是为了看看另一面是否写有元音。但是，既然规则没有说明写有偶数2的卡片另一面会写有元音，那这张卡片便不能决定规则是真是假。

基于这个实验及其他实验，沃森指出，参与者深受偏见之害，这一偏见后来被称为"证实偏见"(注)："结果的确表明……即便是睿智的成年人，也不容易采用科学的态度来对待新奇的问题。若他们能为自身的解

释拿出确切的证据，则会拥有非凡的毅力，能始终坚持己见。”[注]

在你面前有四张卡片。每张卡片一面写有字母，另一面写有数字。有两张卡片的字母面朝上，两张卡片的数字面朝上。



必须要翻开这四张卡片中的哪一张，才能弄清楚以下有关这四张卡片的规则是真是假：“若卡片的一面写有元音，另一面则写着偶数吗？”

图11.1 标准的沃森选择任务

其实，这个实验虽能说服许多人相信确实存在证实偏见，但却没能直接展示这一偏见。沃森和其许多效仿者都主张，参与者会试图“证实”规则而非证伪，其实他们的主张犯了逻辑性错误。事实上，同样的卡片既能在规则有误时证明其有误，也能在规则正确时证明其正确。譬如，在上述例子中，当翻开元音E或奇数7朝上的卡片时，都能揭露规则有误。如果两张卡片都不能证明规则有误，那么这两张卡片便能一起证明规则正确。参与者相当普遍地选择写有偶数2的卡片，但这张卡片不仅不能证明规则有误，也不能证明规则正确。选择写有偶数2的卡片根本算不上一个证实策略。所以，成功完成该任务的人并没有表现出更大的证伪倾向，没能完成该任务的人也没有表现出更大的证实倾向。

如果人们选择的任何给定卡片都不能揭露自身是否存在证实偏见，那么其选择卡片的方式便能。我们发现，人们在选择卡片时，不是靠理性做出快速的决定的，而是靠对不同卡片关联性的直觉。在做出这一直觉反应后，参与者往往会进行长时间的艰难推理。然而，对这四张卡片中每一张卡片的推理，耗费的精力又是不一样的。研究员锁定了参与者的关注点，证实参与者把所有精力都用来思考与规则相关的卡片，即他们已选择的卡片。[注]此外，显而易见，当参与者需要边想边说时，他们

主要思考支撑其直觉选择[注]的理由。真正的证实偏见表现为：参与者不会找寻选择或不选择任何一张卡片的理由，而会找寻支撑其起初选择的充足理由，不会找寻选择其他卡片的理由，也不会找寻不坚持起初选择的理由。

我们很快便能论证，对证实偏见最好的理解其实是“自我中心偏见”。我们先来看一个小案例，该案例充分显示了人们存在自我中心偏见。[注]

学者迪安娜·库恩（Deanna Kuhn）是研究论证和认知的先驱，她让参与者就各种社会问题发表看法，包括失业、辍学、累犯，然后让他们为自己的看法辩护。几乎所有参与者都很乐意提供支撑自己观点的理由。但当库恩要求他们提供反对自己观点的论据时，只有14%的参与者能一如既往给出理由，大多数参与者头脑中都是一片空白。[注]

基瓦·孔达（Ziva Kunda）让一组参与者相信，性格外向者比性格内向者获得成功的可能性更大。所以在之后的记忆测试中，参与者便发现，与内向行为相比，回忆自身的外向行为会更容易。与此相反，他又让另一组参与者相信，性格内向者比性格外向者获得成功的可能性更大。因此这组参与者发现，与外向行为相比，回忆自身的内向行为会更容易。两组参与者都只是在找寻理由佐证自己的信念，即相信自己的性格可以让自己有所成就。[注]

查尔斯·塔伯（Charles Taber）和弥尔顿·洛奇（Milton Lodge）就各种有争议的话题（如枪支管控和平权行动）给参与者展示了各种论点，并让他们就这些论点列出自己的看法。[注]他们把参与者分成两组，一组由不了解政治问题的人构成，另一组则由很了解政治问题的人构成。不了解政治问题的小组成员都表现出了证实偏见：与反对对方立场所列的看法相比，他们为支撑自己一方立场所列的看法多两倍。不过，丰富的知识储备也不能让参与者免除偏见。很了解政治问题的小组成员同样列

出了许多支撑自己一方立场的看法，而对于对方立场，一个反对看法也没有列出。这就证明，政治知识储备越丰富，证实偏见越强。

两组成员所列清单可以排满好几页纸（甚至是好几个章节或好几本书）。而且，我们从鲍林的例子可以看出，深受证实偏见之害的不只是普通的参与者。天赋异禀、聚精会神、积极主动及思维开放的参与者同样不能避免证实偏见。(注)证实偏见“可能是人类推理相关文献中关于推论错误最知名、最广为接受的概念”(注)，有不少实验在证明这一概念的普遍性与稳健性。正如记者乔恩·罗森（Jon　Ronson）讥讽的那样：“自从我了解了证实偏见后，到处都可以看到它的踪迹。”(注)

对唯理智理论的质疑

心理学家一致认为，证实偏见普遍存在。同时，他们也认为有证实偏见不好。证实偏见是“不理性的”(注)且“阻挠了个人能力的发挥，不能使效用最大化”，(注)是“对极端主义意识形态发挥最关键作用的偏见，也是在小组间及小组内冲突中发挥最关键作用的偏见”。(注)雷蒙德·尼克尔森（Raymond Nickerson）恰当地总结了这一普遍性观点。

> 到目前为止，大多数评论员都把证实偏见视为人类的弱点：一个曾经很普遍但不理性的倾向。提出支撑这一立场的充分理由不难。这一偏见会促使人们产生许多不同错觉，促使迷信得到发展并存活下来，促使人们产生各种负面的心境，包括偏执和抑郁。预言家、占卜者和算命先生，甚至任何坚持未经证实言论的人都会充分利用这一偏见。若人们的世界观不一致，憎恶和冲突会一直存在，而证实偏见则在其中发挥着重要作用。(注)

这个评价确实证据确凿。而且，证实偏见并没有隐藏在人类推理心

理学的深处，而是清晰可见的。即使对人性并不敏锐和愤世嫉俗的观察者也认识到人类有证实偏见。譬如，培根为何要费心提醒人们提防“幼稚地”枚举实例的危险？他深知人们会去证实其信念。

> 一旦采纳了某个观点……人类的理解力便会用尽一切办法来支撑和认可该观点。虽然对方有更多更有力的实例，可这些实例要么遭到他们的忽视或蔑视，要么因有区别而遭到搁置或驳回；人们希望通过这种强有力且不顾一切的决心，让其先前结论的权威性可以不受侵犯。(注)

在达尔文发现自然选择的两百多年前，培根就开始著书立说了，但他观察到的只是人们有证实偏见，以及这一偏见阻碍了人们获得合理信念。在当代，捍卫唯理智理论的人认为，理性是自然选择的结果这一点是理所当然的，而证实偏见的存在则向他们提出了根本性挑战。

尤其是要捍卫唯理智理论的人接受以下三个主张：（1）理性有证实偏见；（2）理性帮助孤独的推理者形成更好的信念、做出更好的决定，而证实偏见则会阻碍理性发挥这一功能；（3）理性的主要功能是形成更好的信念、做出更好的决定。其实，接受以下三个主张也是一样的道理：（1）麋鹿的鹿角很大；（2）鹿角太大反而阻碍了麋鹿躲避天敌；（3）这些巨大鹿角的功能是躲避天敌。

这是我们所能想到的解决这一难题的唯一办法，但这一办法行不通。第一种解决方案是证明，证实偏见是理性不可避免的一个特点。譬如，鸟的骨骼、肌肉和翅膀上的羽毛给其躯干增加了额外重量，这些额外重量便增加了鸟躯体飞翔的难度，而不是降低难度。但是，既然我们既不能让翅膀无重量，也不能让鸟没有翅膀就可飞翔，翅膀的重量这个问题本身就不会对进化论提出疑问。但是为什么证实偏见对理性发挥功能很有必要呢？

第二种解决方案是，声称证实偏见对一个次要功能有好处。适应性特性的某些特点可能阻碍了这一特性发挥主要功能，这一点可以通过“这些特点对次要功能有好处”这一事实得到很好的解释。譬如，海雀的重量与翼面积之比（即“翼载荷”）太低了，不利于实现最理想的飞翔。可是，海雀的翅膀却适应水下推进，这就很好地解释了其翼载荷低的原因。[注]与海雀的低翼载荷不同，证实偏见不仅让理性发挥所谓的主要功能变得有点困难，还与理性的主要功能发生直接冲突。此外，虽然海雀的水下推进可以解释其低翼载荷，但理性没有特别说得通的次要功能可以解释其证实偏见。证实偏见应该是唯理智理论中的一个大难题。

直觉是促成证实偏见的罪魁祸首吗

支持唯理智观念的人知道证实偏见很难对付，他们面临一个选择，即否认证实偏见是真正理性的特点，从而为理性脱罪，归咎于直觉。在第二章中，我们已经简要论述了双重过程理论。其实，人们普遍认可这一理论是因为其似乎对解决理性之谜有所帮助：推理缺陷现在可以归咎于类型一直觉过程，而类型二的真正推理则可以免受指责。据称，直觉会出错，因此理性的功能之一便是纠错。[注]譬如，斯坦诺维奇在其2004年出版的《机器人的反叛》（*The Robot's Rebellion*）一书中，将证实偏见列为直觉的一个认知缺陷。[注]埃文斯也指出，证实偏见源自一个更具普遍性的偏见，即“不思考负面信息、只思考正面信息的偏见”。[注]

这一解决方案把证实偏见归咎于直觉，从进化论来看，没有太大意义。直觉推论机制指导我们的想法及行动；自然选择数亿年来已经造就了某些直觉推论机制。直觉信息的质量对我们的繁衍生息非常重要。总的来说，正如我们所见，某些特定偏见可能还是有好处的，譬如它们不仅能降低认知成本，还能降低出现高代价错误的概率。可是证实偏见一点儿好处也没有。

毋庸置疑，我们在研究动物行为时并没有出现证实偏见。小老鼠（mouse）铆足全力证实其认为周围没有猫的信念；大老鼠（rat）则聚精会神地盯着面包屑，而忽视了其他食物，其举动便证实了其认为面包屑就是最好的食物的信念。小老鼠和大老鼠的许多子孙后代都不会遗传这些基因，因为觅食行为会因环境变化而改变。动物一旦能在别处找到更好的食物，就会立马丢掉已获取的小块食物。[注]人类直觉不会比其他动物的推论差。我们遗传了祖先觅食和躲避天敌的能力，还有其他各种推论策略，这些推论策略都没有受到证实偏见之害。[注]

如果真要做什么的话，正如我们之前就主张的，我们应该期待直觉格外关注证伪的、出人意料的信息。心理学家托马斯·艾伦（Thomas Allen）和同事开展了一项实验，该实验很好地展现了直觉和理性各自具有的不同偏见。[注]研究员给参与者看了某人的照片，该照片上还附了两句话来描述那个人的行为，然后他们要求参与者说出对这个人的印象。某些参与者看到的图片是“戴着黑色发带和一副墨镜的年轻黑人男性”，[注]两句描述分别是：“对女售货员一顿痛骂”和“在拥挤的地铁上把座位让给一位老大爷”。如果你熟悉美国社会年轻黑人男性的刻板印象，会发现第一个描述的设计符合大多数参与者的期待，第二个描述的设计却出人意料。

为了测试直觉在形成印象时发挥的作用，研究员促使半数参与者没能运用理性。他们让参与者去牢记一长串数字，由于这项任务占用了持续进行推理所需要的资源，所以这些参与者只能受到自身直觉的指导。如我们所料，他们更多地关注那个出人意料的描述。与之相反，能进行推理的参与者更多地关注意料之中的描述。直觉旨在收集最有用的信息，而推理旨在证实参与者的刻板印象。

自我中心偏见——用处何在

到目前为止，我们都理所当然地认为前述偏见就是证实偏见，即去证实任何碰巧所持看法的偏见。不过，某些实验清楚地揭示，自我中心偏见不能很好地阐述推理的功能。譬如，我们早就发现，参与者很难找到反对其赞成理论的论据。但是，当研究人员要求参与者对其反对观点进行推理时，他们很容易就找到了反对论据。[注]

这些结果——以及许多其他结果[注]——都表明，人们对证实没有普遍性偏爱。他们觉得，总体而言，困难的不是寻找反例或反对论据，而是发现他们自己的观点遭到了质疑。推理不会盲目去证实任何与其相关的信念，而会在我们找寻支撑自身观点的理由及反对某观点的理由时发挥系统性作用。推理总是为我们自己服务的。因此，最好称为自我中心偏见，而不是证实偏见。[注]

既然这个问题已经说清楚了，接下来就扼要重述一下吧。许多证据表明，推理有自我中心偏见。理性很少质疑推理者的直觉，因此不太可能会去纠正推理者可能产生的错误直觉。这恰好与你期待的机制相反，你期待它通过独立推理改善信念。目前，我们还不能运用唯理智理论明确解释自我中心偏见。

强有力且具有普遍性的偏见不可能只是一个错误、一件坏事。相反，它更可能是一个有用的特点。自我中心偏见有何目的呢？沃森把孤独的推理者类比为科学家，这一类比不太合适，而另一个类比却能让我们比较容易地理解这一神秘的偏见。

西塞罗（Cicero）可能是西方文明社会中最杰出的律师，上自罗马的元老院，下至中世纪及现代大学的讲堂，都深受其修辞相关论著的影响。下面是他在《论创造力》（*De Inventione*）一书中建议演说家结束演讲的方式。

> 做到下面两点会很有用：第一，略读你本人单独运用的论据，

同时（仍需更多技巧的事情是）把反对论据与你本人的论据联合起来看；第二，展示你彻底否决反对论据的方式。这样，通过简单对照，便能唤起听众的回忆，既回忆起你的引证，又回忆起你的非难。(注)

换句话说，就是当你想说服某人时，只给出支撑你个人立场的论据，或只给出你所反对立场的反对论据。西塞罗坦率地支持自我中心偏见，而且这完全讲得通。如果某个律师开始据理反对其委托人或为对方辩护，他很快就会失业。(注)

律师类比的例子使人联想起一个语境，其中最重要的是说服力，而且自我中心偏见的意义也显而易见：当捍卫某个观点时，有自我中心偏见是好事。(注)它是特点，而不是错误。这符合互动理论的预言。如果在制造理由时，推理的功能是为自身行动辩护或说服他人，那么这项推理应该存在自我中心偏见。

做自己的律师

如果自我中心偏见在理性中根深蒂固，那么为何西塞罗要提醒我们要有自我中心偏见呢？虽然“理性是律师”的类比清晰易懂，与更经典的类比（把理性类比为科学家）相对照后更是如此，但是我们却不能对这一类比进行更深层次的阐述。譬如，与经典的推理者不同，律师通常要为其本人并不认可的立场辩护。他们的“立场”就是雇用其来辩护的委托人的立场。我们很难找到论据支撑自己并不赞成的立场，或自己反对的立场。要做到这一点，我们需要学习技巧并加强训练。我们可以成为律师，但前提是我们是在捍卫自己认可的信念及决定。

我们不能对律师类比进行更深层次的阐述，因为还存在其他理由。

律师在准备辩词时，除了理性，许多其他的认知机制也在发挥作用。当他们有意压制对其委托人不利的论据，或利用这种论据来推测对方后面的辩词时，除了依靠普通推理，还依靠策略规划。在陈述的理由中表现这种事先预演的一致性，推理新手还做不到这一点。此外，作为律师，无论如何都要一心一意为自己的委托人辩护。然而，普通推理者虽然有自我中心偏见，但当其发现或偶然碰到反对自己观点的论据时，若认真对待这些反对论据，甚至为此改变自己的看法，可能对其而言最有利。

最重要的是，我们只有在拿出论据时才会用到律师的类比。在法庭审判中，每个当事人都是一个认真设定过的角色。稍微简化一下，就是让律师拿出论据，让法官和陪审团评估这些论据。要评估论据，法官和陪审团也要依靠理性，但他们与律师不同，不应该存有偏见。在第十二章中，我们会看到怎样评估论据才符合理想化的公正法官形象。

在对抗性审判中，敌对双方被困在零和博弈里：一方赢，另一方就得输。虽然零和博弈强调了自我中心偏见的实用性，但也会不必要地与竞争环境联系起来。事实上，即使人们在寻找好的解决方案时有共同的利害关系，并因此参与到了正和博弈中，拥有自我中心偏见仍会是持续进行博弈的最好方式。

假设两个工程师必须要想出一座桥的最佳设计方案。无论最终选择谁的设计方案，他们都会一起监督施工，因为其目的只是把桥建好。艾拉（Ella）想建吊桥，迪克（Dick）想建悬臂桥。要解决这个矛盾，他俩需要认真考量对这两种方案正反两方面的观点，然后衡量两个方案的可行性，最后给出评分。接下来他们只需要计算出平均分就可以，不需要进行论述，但需要做大量研究工作。

他俩可以交替为各自的方案进行案例分析。艾拉收集赞成建吊桥、反对建悬臂桥的观点，迪克收集赞成建悬臂桥、反对建吊桥的观点。接着他俩基于对双方论据的聆听与评估，再商榷哪个方案更好。这一例子表明，与自己找寻论据相比，评估他人提供的论据会更容易，因为这样

做既减少了我们的工作量，又能让我们获得同样的结果。艾拉和迪克两人只需分别找到一半论据即可，然后再仔细审查每个方案正反两方面的观点。

自我中心偏见没有把论证变成纯竞争性尝试。论证是一种沟通形式，且通常协同进行。自我中心偏见最好是成为分配认知劳动力的方式。在第十二章中，我们会看到，有一个类似的驱动力在发挥作用，其所用方式与理性评估论据的方式相同，但与找寻论据的方式不同。

---

1. “Linus C.Pauling Dies at 93; Chemist and Voice for Peace”, New York Times, 1994.8.21，网址：http://www.nytimes.com/learning/general/onthisday/bday/0228.html。
2. Watson 1997, Kindle location 395.
3. Maurice Wilkins, BBC radio 4 interview, 1997, Oregon State University Libraries，网址：http://oregondigital.org/cdm4/item_viewer. php?CISOROOT=/dna&CISOPTR=166&CISOBOX=1&REC=1。
4. “Linus Pauling Rebuts New Mayo Study on Vitamin C”,1979.9.28，p.2，网址：https://profiles.nlm.nih.gov/ps/retrieve/Resource Metadata/MMBBKK（也可参见：Collins & Pinch 2005）。
5. Creagan et al. 1979.
6. Moertel et al. 1985.
7. Collins & Pinch 2005.
8. Pauling & Herman 1989.
9. Pauling & Herman 1989，p.6837。
10. Pauling & Herman 1989，p.6835。
11. Pauling & Herman 1989，p.6835; 着重进行贬损性“描述”。
12. Mansel Davies,“Obituary: Professor Linus Pauling”,Independent, 1994.8.21，网址：http://www.independent.co.uk/news/people/obituary-professor-linus-pauling-1 377923. html。
13. 《牛津英语词典》，网址：https://en.oxforddictionaries.com/definition/bias。
14. Tversky & Kahneman 1973.
15. 例如，Gigerenzer 2007。

16. Funder 1987; Haselton & Buss 2000.

17. Ings & Chittka 2008.

18. Liszkowski et al. 2004.

19. Crombie 1971, pp.84ff.欲从另一个角度解读格罗斯泰斯特及其证伪理念，可参见：Serene 1979。

20. 两个引用都出自Bacon 1620, bk.1, p.105。Crombie 1971将格罗斯泰斯特与培根联系起来，Urbach 1982指出培根与波普尔就这一问题的共性。

21. Popper 1963. p.6.

22. Mynatt, Doherty, & Tweney 1977.

23. Wason 1968, p.142.

24. Ball et al. 2003; 欲了解相关的方法论依据，可参见：Evans 1996。

25. Lucas & Ball 2005.

26. Nickerson 1998, p.175.

27. Kuhn 1991，有160个参与者，通过第142页的表5.4估算而得。

28. Sanitioso, Kunda & Fong 1990.

29. Taber & Lodge 2006.

30. Johnson-Laird & Byrne 2002; Stanovich & West 2008; Stanovich, West & Toplak 2013.

31. Evans 1989, p.41;但是并非所有的演示都有相同的说服力，可参见：Mercier 2016b。

32. “Julie Burchill & Psychopaths”,n.d., 网址：http://jonathanron-son.tumblr.com/post/43312919747/julie-burchill-and-psychopaths。

33. Evans & Over 1996, p.20.

34. Stanovich 2004, p.134.

35. Lilienfeld, Ammirati, & Landfield 2009, p.391.

36. Nickerson 1998, p.205.

37. Bacon 1620.

38. Thaxter et al. 2010.

39. Evans 2007; Kahneman 2003b; Stanovich 2004.

40. Stanovich 2004, p.134.

41. Evans 1989, p.42.

42. Pyke 1984.

43. 欲搜索相关信息，可参见：Hill et al. 1987。

44. Allen et al. 2009.

45. Allen et al. 2009，p.1083。

46. Shaw 1996, p.80.

47. Mercier & Sperber 2011.

48. 该术语曾在Baron 1995 & Perkins 1989中以稍微不同的含义出现过。

49. Cicero,De Inventione, bk.1, chap.52, trans.C.D.Yonge, 网址：http://classicpersuasion.org/pw/cicero/dnv1-4.htm#97。

50. 的确，要是替你辩护的律师的自我中心偏见前后不一致，就会招致上诉，因为该辩护律师“让诉讼程序渐渐不能正常运作”，Strickland v.Washington, 466 U.S.668 (1984), 网址：http://supreme.justia.com/cases/federal/us/466/668/case.html（辩护律师的无效辩护）。

51. 欲了解先前类比的一些例子，可参见：Knobe 2010; Tetlock 2002。

第十二章
质量控制：如何评估论据

在第十一章中，我们看到，西塞罗的建议是彻底倒向一边，由此看来，他不过是个“大男孩”，这样的建议“粗略且不完整”。多年以后，他作为罗马演说家，“已从众多他为之申辩的案例中积累了丰富的经验”，注因此，其看法发生了很大改变。

> 一个人若能效仿亚里士多德，对各领域问题都能发表看法，而且兼顾到问题的两面，另外还从亚里士多德的训令中学会就任何话题发表两场自相矛盾的演说……同时运用修辞技巧，并在演讲中反复练习，做到这些，他就会成为真挚诚恳、完美无瑕、独一无二的演说家。注

他作为久经参议院辩论场的老手，轻而易举便能堆砌论据来支撑一方立场。演说家要做出漂亮的陈述，必须结合修辞技巧与预测能力，并考虑到可能出现的反对论据。尽管演说者仍然只会陈述支撑自身立场的论据，但还是应该机灵点，不能被自我中心偏见蒙蔽。

我们认为，若发言者是为自身观点及行为辩护，或为说服他人，自我中心偏见便合乎情理。但若理性允许我们去找寻反对自身立场的论据（哪怕只是为了辩驳这些论据），我们不是能拿出更合理的理由、形成更坚定的信念吗？身经百战的演说家会花时间打磨其论据的内容、风格和呈现形式。无论是西塞罗、昆体良（Quintilian），还是当代的演讲稿撰写者、公关专家，都耗费了大量精力来设计有说服力的论据。若真如

我们声称的那样，推理进化是为了改变他人看法，它不是应该去找寻这种精心设计的论据吗？

下面便是我们的互动理论面临的潜在问题：大量的证据、对互联网评论的基本了解表明，人们远非西塞罗般的人。在第十一章中，我们看到了迪安娜·库恩的论证研究，其研究表明大多数人找寻反对论据时都有困难。这一研究还表明，人们找寻的支撑自身立场的论据通常也相当无力。当实验人员问参与者如何看待学生学业不及格时，该参与者回答，营养不良是主要原因。以下是两人谈话的一段摘录。

> 实验人员：若你要试图说服他人相信你的观点正确，会拿出什么证据来证明呢？
>
> 参与者：学生在学校获得的分数。通过他们在学校获得的分数来证明……
>
> 实验人员：这会证明什么么呢？
>
> 参与者：证明他们的身体里缺少某种营养。孩子学业不及格是因为他们身体里缺少某种营养。[注]

正如库恩指出的那样，这位参与者“清楚地认为，学生学业不及格，这个现象的存在本身就可以作为充足的证据，而这也就是该参与者提出的理由”。[注]其他参与者也给出了解释，但也只是其看法的重述而已。大多数参与者都想不出证明其观点的证据。库恩对未受训参与者的辩论技巧评价极低，其他杰出的心理学家也赞同其观点，如理查德·尼斯贝特和李·罗斯（Lee Ross）指出，人们碰巧想到一个理由后通常就止步不前了，满足于现状。[注]另外，大卫·珀金斯（David Perkins）还断言，许多论据只有“肤浅的意义”。[注]

显而易见，我们推理时并不总能注意到有说服力的决定性论据：寻

找合适的理由成本很高且很费时。应在理由的品质与找寻它们所耗费的心力之间求得平衡。但是，这些心理学家声称，我们不应该找寻虽不完美但十分合适的理由，而应该找寻十分肤浅且无力的理由。

这是唯理智理论面临的又一个问题。若理性的功能是改善孤独推理者的信念及决定，那么它最好给出合理的理由。事实上，人们为自身理由设下的标准相当宽松。因此，人们通常很容易找到理由来支撑自身直觉（自我中心偏见），但是不能确信其找寻理由的合理性。理性总体而言还是没能纠正错误的直觉，这一点没什么可惊讶的。

若理性旨在改变他人想法，又如何呢？如果人们表现得像律师那样，花费时间和精力来预测可能出现的反对论据，并找到更好的论据来反驳，这样不是更有说服力吗？理性通过进化，不仅可以用在法庭上，还可以用在非正式语境中。这样不仅大大降低了日常讨论以失败告终的风险，其讨论进程也会有很大不同。讨论是互动性的：不同于长篇大论且精心设计的辩词，人们在讨论中所交换的论据通常都短小精悍。这种你来我往的讨论促使人们形成合理的论断，还不用费多大气力。

互动性论证更容易

社会语言学家强调对话双方如何互帮互助以实现有效的沟通，如给发言者不断提供反馈：“嗯嗯”“是的”、点头等。人类学家、语言学家史蒂文·莱文森（Steven Levinson）沿用这一传统，主张人类天生就有“互动引擎”。[注]我们的交际能力可以适应交际语境，从而自然地发挥功能。例如，当我们想要指代某人时，通常面临各种各样的选择，如“凯瑟琳·特克女士”（Ms.Catherine Turk）、“特克女士”（Ms.Turk）、“凯特”（Kate）、“审计部门主管”等。在给定的语境中，排除其他因素，选择哪一个称呼最合适还取决于对话双方对这个人的熟悉程度。幸好，若称呼用错，很容易就能予以更正。[注]

迈克尔（Michael）：我午饭是和凯特一起吃的。

罗博：谁？

迈克尔：凯特·特克。

罗博：我刚到这儿，我不认识她。

迈克尔：抱歉。她是审计部门的主管。

罗博：啊，原来如此。

迈克尔刚开始用了在美国最稀松平常的称呼方式，即用“名”来称呼熟人。罗博的反应促使他更正自己的表述，来确保有效沟通。

要是不利用互动搭档的反馈，需要思考良久才能决定哪个形式最合适。

迈克尔思考的是：这里每个人都认识凯特。但是罗博刚来。不过，凯特总是努力向每个新员工做自我介绍。罗博是两天前来的，所以他应该认识凯特。但我前两天都没看见凯特，所以她或许生病了，没有来。我最好详细介绍一下她，以防万一。

人们先采用看似最合适的称呼，然后若有必要，便根据反馈进行更正，这种做法是最省时省力的解决方案。沟通者认为，使用人们可以理解的最精简称呼形式，即“懒惰之举”，有时也是明智之选。注至少有两个理由可以解释为何对话方的反馈有用。其一，对话方的理解才是最重要的，让他说出他是否已理解很容易。其二，对话方不仅可以表明自己没有领会说话者的意思，还能积极地引导说话者，正如罗博告诉迈克尔他为何不认识凯特一样。

推理者也面临类似的挑战，即要找出传达信息的最好方式。同样，经过精心推测，他们在任何情况下都能找到最合适的论据。这一策略是

时间和精力密集型的，但并非万无一失，下面的交谈会进行阐明。

夏洛克·福尔摩斯在咖啡店遇到好友华生。

福尔摩斯：亲爱的华生，你有那么迷人的女伴在场，我没敢打扰！我已经观察你俩好久了，我强烈建议你再约一下这位女士！她跟你非常般配。虽然她比你年龄小，但也差得不多。我知道，你喜欢深褐色头发的女人，而她是金发，但无疑应该忽视这一差异。我还观察到，你们在谈及私事时，没有任何亲密的肢体接触，但我相信，你们很快就会有的。华生，我敢保证，这个女人最适合你！

华生：她是我妹妹。

无疑，柯南·道尔（Conan Doyle）总能操控情境，这样的话，他笔下的福尔摩斯就不会犯这种失言错误。福尔摩斯这个虚构人物，说他像个小男孩更现实，因为他没猜到华生迷人的女伴竟是他妹妹。如果他在惊呼“你女伴真迷人！”后停顿一下，华生就会告诉他这是他妹妹，这样的话，也能避免进一步的尴尬。

即便是在形式简单的论证中，反馈也发挥了重要的作用。拿下面这段司空见惯的对话举个例子。

伊莲娜（Hélène）：我们应该去伊佐美（Isami）吃饭。这个餐厅不错。

马乔里（Marjorie）：我不知道该去哪儿。我上周去过日本餐厅了。

伊莲娜：但这个十分新颖。

伊佐美新颖独特、价格实惠、鱼片肉质鲜美、座无虚席，综合这些因素，会是个好去处，但是伊莲娜一开始并没有将这些理由全部列出

来。相反，她只给出了一般的概括性评价：“这个餐厅不错。”这个论据并没有说服马乔里，但她没有简单回复“不去”，而是给出了自己的理由：“我上周去过日本餐厅了。”多亏了马乔里的反馈，伊莲娜才得以调整其下一个论据，指出伊佐美的品质，因为这样最有可能改变马乔里的想法。换个语境，对话会变成以下这样。

伊莲娜：我们应该去伊佐美吃饭。这个餐厅不错。

马乔里：我不知道去哪儿。我现在没有太多钱，日本餐厅很贵。

伊莲娜：但这个餐厅很实惠。

虽然伊莲娜一开始可以给出许多论据，但只有一部分更能说服马乔里。伊莲娜本可以尝试预测哪个论据说服力最强，但那样做会花些工夫。而且鉴于她可以利用所有相关信息，譬如马乔里上周去日本餐厅吃过饭，又譬如她囊中羞涩，所以她不大可能对多个论据的说服力进行预测。

不可否认，当机立断有时至关重要。当法国启蒙运动领袖伏尔泰（Voltaire）要被英国暴民处以私刑时，他必须立即说服这些人相信他真的是一个亲英者才能自救（显然他做到了）。幸好，只有极少数论述会这般紧急。大多数论述都是，该谁洗碗、是否应邀请琼斯（Jones）一家来参加晚宴，或电影院当前上映的哪部电影最好看。就算在平淡无奇的探讨中没能迅速占上风，也几乎不用付出任何代价。即便赌注更高时，如顾客是否要买车、警察是否要接受你的陈述、你们一家是搬到新加坡还是待在中国香港，对这些问题，没能用第一个理由说服对方一般也不要紧，还有更多理由可以尝试。

倘若拿出一连串无力理由的代价是让说话者看起来愚笨不堪，那么，即便推理者比较随性，在给出理由时，也应该适当提高质量。不

过，显而易见，我们的互动理论并不认为人类天生就应该是西塞罗般的人，能像他一样编织复杂的论据，能自发地设想可能会遭到的反驳。理性应该充分利用对话的互动性质，利用对话方的反馈来完善正当理由及论据。

完善理由

先前的实验并不是在典型的对话语境下开展的，这就促使心理学家公开谴责参与者提出的理由质量低下。若普通的对话方没有被说话者的理由说服，就会给出反对论据，从而促使说话者拿出更有力的理由。与此同时，实验人员会保持中立，会促使参与者给出更多论据，却不会反驳。如果理性通过进化，会在互动中发挥其功能，则只有在出现了同样强有力的反驳时，强有力的论据才会出现。为评估人们在真正的辩论环境中构造论据的能力，劳伦·雷斯尼克（Lauren Resnick）及其同事让学生分小组讨论有争议性的话题。参与者每分钟都会沟通好几个论据，提出新观点以供讨论，然后评论各自的建议及论据。在这样你来我往的讨论中，参与者表现得都很好，“推理的一致性也让研究员印象深刻。参与者……似乎在构造复杂的论据并破坏对方的论证结构。人们似乎能识别这些结构并有效抨击个人成分及整个论据”。[注]迪安娜·库恩和其同事在更为量化的研究中得出了类似的结论，[注]他们发现：当学生之间就死刑问题进行一场辩论后，便能拿出有力得多的论据。他们阐明了前提与结论之间的联系，补充证据支撑自身看法，并依赖多种论据类型。

在第十一章末，我们看到，每个个体都能找到支撑自己立场的论据并对支撑对方立场的论据进行评估，从而促使自我中心偏见成为分配认知劳动力的有效方式。我们刚描述的这个过程是认知劳动力的又一种巧妙分配。参与者没有搜寻代价高且可能无望的决定性论据，而是依靠对话方的反馈，针对具体的反对意见，对自己的论据进行调整。

谬论？什么谬论？

人们对自己的理由实行低质量控制，轻易就满足于相对无力的正当理由及论据。从互动论视角来看，鉴于理性通过进化以在互动语境下运作，人们的表现便很容易解释了。正如预料的那样，与独自推理相比，在对话性互动中，人们给出的论据更合理、更尖锐。

对他人理由的质量控制又如何呢？互动理论做出了非常不同的预测。如果你在评估自己的理由时松懈了，那么在评估他人理由时就会有更高要求。不然的话，我们就会把最愚蠢的借口当作合理的辩护，还会把最明显的谬论当作合理的论据。我们很容易就会被操控。普遍观念认为，人们很容易被诡辩所骗，而这一预测则与之相悖。普遍观念是错误的，部分原因是它依赖评估论据的错误标准。实验数据表明，若所用标准更合乎情理，人们往往会在应该认可时去认可所给理由，仅仅是在应该这样做时。

不论是诉诸公众（ad populum）的谬论，还是诉诸人身攻击（ad hominem）的谬论，辨别论据好坏的普遍性标准都在于这些论据能否被归为非形式谬论。关于这些谬论的清单在古典时期就已经列出来了，如今在网络上也很容易就能找到。问题是，对于该清单中的几乎所有谬论类型，都能找到反例，且都是符合标准的论据形式，虽然宣称这些反例靠不住，但在现实生活中都是可接受的、合乎情理的论据，连理性的受众都可以被说服。㊟

下面是诉诸伪善（tu quoque，拉丁语，用作反驳）形式的谬论。

耀西（Yoshi）：既然你要开车，就不应该喝酒！

真希子（Makiko）：一个月前你难道没有因酒驾被抓？

真希子反对耀西建议所给的理由只是荒谬的论据。毕竟，耀西没能遵循自己所给建议不能代表其建议是错误的。另外，如果你怀疑说话者不遵循自身建议是有合理的理由的，那么诉诸伪善形式的论据会变得非常合情合理。

耀西：你不应该吃伊莲娜婶婶买给我们的这些巧克力，它们不太好吃！

真希子：这一盒所剩无几了，难道不是你吃的吗？

在这种情况下，真希子的异议是对耀西建议可信度的合理质疑。

下面是诉诸无知（ad ignorantiam）形式的谬论（认为某断言正确就因为没人认为它是错的）。

警察：我敢肯定石井（Ishii）是间谍。我找不到证据证明他不是间谍。

人们普遍认为，警察给出的论据是荒谬的，因为不知道某命题有误一般不能作为说明该命题正确的合理论据。但是，在某些情况下，这种论据确实非常合理。

警察：我敢肯定石井是遵纪守法的好市民。我找不到证据证明他不是。

我们还可以举出更多例子，说明的都是同一种情况：除去少许例外情况，诉诸伪善形式的谬论通常都是荒谬的，诉诸无知形式的谬论亦是如此。实际上，并非所有谬论都是荒谬的，还是有少许不荒谬的谬论。若对谬论进行更加准确的界定，人们通常会含蓄地承认这一点。诉诸伪

善形式的谬论是对“对话方没有坚持自身看法”这个事实的不合理利用，诉诸无知形式的谬论是对“无人知晓经检验的主张有误”这个事实的不合理利用，等等。

这种定义谬论的方式让我们可以捏造出各种新类型：珠穆朗玛峰谬论（在论据中不合理地与珠穆朗玛峰形成对照）、鸡汤谬论（在论据中不合理地利用有关鸡汤的事实）等。要知道，政治哲学家列奥·施特劳斯（Leo Strauss）还捏造了希特勒归谬法（reductio ad Hitlerum）谬论，表现为在反对某观点时将之不合理地比拟为阿道夫·希特勒（Adolf Hitler）会持有的观点。与荒谬地参照珠穆朗玛峰相比，荒谬地参照希特勒要频繁得多，所以为何不享受给这种参照贴上专用名标签的乐趣呢？

心理学家乌尔丽克·哈恩（Ulrike Hahn）、麦克·奥克斯福德（Mike Oaksford）及其同事抨击了标准非形式谬论的观点。[注]他们详述了某些变量，此举应该会使某些论据相比其他论据更有说服力。譬如，在诉诸无知谬论中，主变量之一便是：若要找寻正面证据，成功的可能性有多大？间谍案例中给出的论据很无力，因为即便是警察，也很难找到证据证明某人不是间谍，如果间谍那么容易被识别，他们就不会潜伏那么久了。与之相比，警察要找寻证据证明某人是不良青少年，若事实真的如此，证据很容易就能找到，若没能找到，那便证明这个人不是不良青少年。此外，幸运的是，间谍比良好市民更为罕见，这便进一步削弱了间谍案例中论据的可信度。

哈恩和奥克斯福德并不只是纸上谈兵，他们还通过操纵相关变量、让人评估给出的论据检验了他们的假设。通过观察大量实例，他们发现人们用理性的方式评估论据，不会轻易受到真正荒谬的论据之害。

譬如，下面是他们让参与者评估的其中一个诉诸无知形式的论据：

这个药物很可能并没有副作用，因为五项严格控制的、大规模的临床试验都没能发现其有任何副作用。

人们看到这个论据时，认为十分有说服力。若论据发生变化，则参与者的反应也会相应地改变。譬如，若只有两个试验没有揭示任何副作用，参与者就会有所动摇。不管论据类型如何，都能得到相似的结果。总之，有证据表明，若他人给出的论据是真的荒谬，人们则相当擅长否决它们。㊟

把自己的理由当作他人的理由进行评估

我们回顾的这些实验表明，人们拿出自己的理由与评估他人的理由的态度之间存在不对称性，他们对自己的理由的质量控制相对宽松，而对他人理由的质量控制则要严格得多。在伊曼纽尔·特鲁奇（Emmanuel Trouche）、彼得·约翰逊（Petter Johansson）和拉斯·霍尔（Lars Hall）的协助下，雨果开展了一项棘手的实验，旨在使这种不对称性尽可能明显。这就需要人们把自身论据当作他人论据进行评估。㊟

在该实验的第一阶段，参与者解决了五个与果蔬店售卖果蔬相关的简单推理问题。譬如，他们可能听说，果蔬店里的“苹果都不是有机的”。根据这个信息，他们必须尽快得出直觉结论，在“有些水果不是有机的”和“我们不敢肯定这个果蔬店里的水果是否有机”等推论中做出选择。譬如，叫拉万（Rawan）的参与者就选择相信“有些水果不是有机的”这个推论。

在实验的第二阶段，实验人员要求参与者给出支撑其五个直觉回答的理由，这五个直觉回答是在实验第一阶段做出的。给出理由后，只要他们愿意，还可以修改起初的答案。鉴于对理由制造已经有所了解，我

们此刻不应抱有太多期待。此外，大多数参与者给出支撑自身直觉的理由时竟没有仔细检查其合理性，也没有修改起初的选择。最终只有14%的参与者改变了看法，改变的结果可能更好，也可能更糟。拉万就是没有改变看法的参与者之一。关于有机水果问题，拉万给出了以下理由来支撑自己的答案：“因为苹果都不是有机的，苹果是一种水果，所以我们可以推断，果蔬店里有些水果不是有机的。”

在实验的第三阶段，实验人员给每个参与者提了五个同样的问题，并提醒他们之前给出的答案是什么。接着，实验人员告诉他们，昨天有个参与者给出了不同的答案，并给他们看这个参与者的答案和论据。基于这一论据，他们可以改变自身看法，认可该参与者的答案，也可以坚持起初的答案。

我们利用其中某个问题“捉弄”了参与者，告诉他们现在与原先的答案不一样。譬如，我们告诉拉万，她之前的答案是“我们不敢肯定这个果蔬店里的水果是否有机”。此外，我们还告诉她其他人的选择是“有些水果不是有机的”这个推论，理由是“因为苹果都不是有机的，苹果是一种水果，所以我们可以推断，果蔬店里有些水果不是有机的”（其实这就是拉万本人的选择和论据）。

我们希望用这个完美的圈套来检验理由制造与理由评估之间存在的这种不对称性：实验人员要求参与者评估某个论据，这一论据是参与者几分钟前给出的，而其本人却没有意识到。这一方法果真有效。包括拉万在内的约半数参与者都没有注意到，他们被骗了，把自己的理由当作了别人的。

我们成功地误导了这些参与者，他们最起码应该会认同自己前一刻给出的论据吧？当然不会。即便他们认为自己的论据足够充分，但当他们误以为这是他人的论据时，会用更加批判性的态度看待它，而且大多数时候都认为该论据不够充分。令人欣慰的是，与自己的合理理由相比，自己的不合理理由更可能遭到参与者的否决。拉万起初便给出了合

理的理由，当她误以为这一理由是别人给出的时，还是同样认为其具备说服力。

认可合理的论据

到目前为止，我们已经着重指出了理由评估的一个重要特点：在否决没有说服力的理由时，要足够严苛。但是在认可合理的理由时，严苛也很重要。在第十章中，我们主张，推理对沟通者及受众都有作用。若沟通者想仅凭自己的威信做出受众不可能认可的断言，推理则产生受众会评估并认可的论据，接着促使受众认可该断言。对受众而言，推理是认知警觉工具。受众通过推理来评估沟通者给出的论据，若该断言的论据不充分的话，则拒绝接受该断言；若该断言的论据很充分的话，则认可该断言。当然，认知警觉不仅让人拒绝接受可疑信息，还让人获得有用信息。因此，若他人给出的理由足够充分，则我们要适时改变自身看法。

为检验这一预测，在伊曼纽尔·特鲁奇和邵荆（Jing Shao）的协助下，我们用下面这个问题——被称为“保罗（Paul）和琳达（Linda）”问题——开展了一系列实验。(注)

> 保罗暧昧地看着琳达，琳达也暧昧地看着约翰。保罗已婚，但是约翰未婚。现在的问题是：一个已婚人士正暧昧地看着一个未婚人士吗？

回答有三种，分别为“是”“不是”“不确定”。好好想想这个问题，真的很有意思。

大多数参与者回答“不确定”，认为要给出确切的答案必须要知道琳

达是否已婚。但是细想一下下面这个论据。

> 琳达要么已婚，要么未婚。如果她已婚，那么她正暧昧地看着约翰这个未婚人士，所以应回答“是”。如果她未婚，那么保罗这个已婚人士正暧昧地看着她，所以还应回答“是”。因此，无论是哪种情况，琳达已婚还是未婚，回答总是“是”。

若你与我们的参与者（网上招募的美国人和中国人）无论在哪方面都相像，则很可能会认可这一论据。当参与者看到该论据时，半数以上的人立即改变了看法。相比之下，若你独自弄清楚了该问题，还听说“答案是‘不确定’，因为我们不知道琳达是否已婚”，则绝不会改变自己的看法。人们评估此类论据的方式十分稳健。

当我们给参与者看支撑正确答案的论据时，没有告诉他们这是我们（作为实验人员）给出的论据，而是告诉他们这是先前实验中别的参与者给出的。所以，他们没有理由不相信自己，而选择去相信他人。我们还添油加醋地对某些参与者说，给出这个论据的个体真的不擅长解答该类任务。此外，我们告诉其他参与者，若他们回答错了，给出这个论据的个体就会拿到一笔钱。因此，参与者对他人给出的论据抱有期待，尽管他们觉得这个人要么太蠢，要么存心欺骗他们。结果，当参与者看到这个论据时，大多数人都接受了。虽然参与者的确说过丝毫不相信给出该论据的人，但是这种信任匮乏对其认可该论据几乎没有影响。

我们也让某些参与者认真思考这个问题，并佐证其答案。少数参与者能彻底搞清楚该问题。大多数参与者还是理解错了，他们觉得自己认真思考过了，所以给出的答案不会有问题，但其实是有问题的。大多数参与者说，他们“极其自信”，甚至说“没有比这更自信的事了”。但见到上面这个论据时，尽管他们很自信，还是会改变自己的看法，而且与心存疑虑的参与者相比，他们改变自己的看法的可能性有过之而无不及。即便他们本可以郑重声明该推论有误，但当他们看到这个论据时，还是

受到了影响。

理性两面论

在本章及第十一章中，我们研究了推理制造论据及评估他人论据的方式。在表12.1中我们对此进行了归纳总结。

表中的“理由制造”一行对唯理智理论真的不利。若人们独自推理，主要会拿出支撑自己决定或预想的理由，还不用费事去证实其说服力。在第十三章中，我们会看到，这是应对祸患的秘诀：单独运用理性不仅不可能纠正错误直觉，甚至还会使事情变得更糟。

事实是，人们擅于评估他人的理由，这对唯理智理论形成了致命一击。这意味着，人们能进行客观推理，从而拒绝无力论据、接受有力论据，也意味着，他们在制造理由时并没有用到这些技巧。显而易见，理由制造的缺陷不是认知失灵，而是认知特点。

表12.1符合唯理智理论的预测。如该理论预测的那样，人们制造理由：找寻理由支撑自己的立场（若其目标是改变他人看法，则此举适宜），从容易找到的理由入手，而不是起初就给出最有力的理由，这样就充分利用了对话环境下的反馈。人们也按照预想去评估他人理由，拒绝无力的理由，认可足够有力的理由，即便要修改自己强有力的信念，抑或倾听不值得信赖的沟通者。

表12.1 理性两面论

	偏见	质量控制
理由制造	有偏见的：人们主要制造理由支撑自身立场	懒惰的：对于自身理由不会过分苛刻
对他人理由的评估	没有偏见的：若理由足够有力，即便其富有挑战性，人们也会接受	高要求的：只有理由足够充分，人们才会深信不疑

若我们从互动论视角看问题，在论据制造过程中通常被视为瑕疵的特点，便成了分配认知劳动力的巧妙方式。推理过程中最困难的任务是找寻合理的理由，而自我中心偏见和适当的懒惰却使这一任务变得容易得多。自我中心偏见促使推理者只关注问题的一个方面，而不必独自弄清楚要怎样采纳每个人的看法。一旦理性找到了合适的理由，就会受到懒惰心理的影响，进而停止找寻更合适的理由。若对话方没有被说服，就会去找寻反对论据，进而促使发言者制造更多有说服力的理由。理由沟通借助偏见和懒惰的优势，提供了巧妙且高效的解决分歧的方式。

---

1. Cicero,De Oratore, bk.1, chap.2, p.7, trans.J.S.Watson, 网址：http://archive.org/details/ciceroonoratoryaoociceuoft。
2. Cicero,De Oratore, bk.1, chap.2, p.7, trans.J.S.Watson, 网址：http://archive.org/details/ciceroonoratoryaoociceuoft，bk.3, chap.21, p.214。
3. Kuhn 1991, p.87.
4. Kuhn 1991, p.87.
5. Nisbett & Ross 1980, p.119.
6. Perkins 1985, p.568.
7. Levinson 2006;也可参见：Enfield 2013。
8. Levinson 2005为之提供了些许例子。
9. Dingemanse & Enfield 2015; Dingemanse et al. 2015.
10. Resnick et al. 1993, pp.362-363.

11. Kuhn, Shaw, & Felton 1997.

12. Boudry, Paglieri & Pigliucci 2015.

13. Hahn & Hornikx 2016; Hahn & Oaksford 2007; Oaksford & Hahn 2004.

14. 欲了解更多证明人们善于评估论据的证据，至少是在受到激励的情况下，可参见有关说服力与态度变化的文献。Petty & Wegener 1998给出了较好的评论。

15. Trouche et al. 2016.

16. Trouche, Shao & Mercier这样主张；也可参见：Trouche, Sander & Mercier 2014。

第十三章
理性之阴暗面

柯南·道尔的小说《巴斯克维尔的猎犬》（*The Hound of the Baskervilles*）开篇是莫蒂默博士（Dr. Mortimer）想雇用夏洛克·福尔摩斯为他工作：

> “我是来找你的，福尔摩斯先生。因为……我突然遇到了一个非常严重、非常离奇的问题。我认为——嗯，一直这么认为——你的破案能力在欧洲可排第二……”
>
> “确有其事，先生！我能问一下是谁荣为第一吗？”福尔摩斯带着一丝严厉的口气问道。
>
> “是个叫贝迪永（Bertillon）的人，他有精准的科学头脑，他破的案子总是十分引人入胜。”
>
> “那么你去找他不是更好吗？”㊟

这个贝迪永先生是谁呢？是个比福尔摩斯更厉害的神探吗？阿方斯·贝迪永是19世纪末最受尊敬的警察，还是世界上最擅长司法鉴定的科学家。他基于人体测量学，发明了一套记录罪犯身份的科学方案，这一方案能精准测量罪犯的各种特点，从左脚的大小到右耳的长度。㊟这一体系中最著名的要素是入案照片，是拍摄被捕嫌疑犯的标准化方式，沿用至今。

鉴于贝迪永在摄影这一新领域的专长，早在1884年10月，法国军方

调查员就向他发出了邀请，任务就是给某极其敏感的文件拍摄一连串的照片，该文件是一封信函，被称为“清单”（bordereau），写信的人是一名德国间谍，在法国做军官。贝迪永同意帮忙。而几年后，这张小纸片让他变成了法国最荒唐可笑的人。

10月13日，贝迪永又被派去执行一项任务，这项任务超出了其专业知识领域。为了确定清单的作者，他必须判断出另一封信的字迹是否与清单的字迹匹配。几小时后，贝迪永推断，两封信很可能是同一个人写的。然而，对领导这次审讯的亨利（Henry）中校和迪帕蒂·德克拉姆（du Paty de Clam）中校来说，证据不够确凿。他们想要一纸证据确凿的起诉状。他们两天后再次拜访了贝迪永，并给了他更多信息：他们有证据证明艾伯特·德雷福斯（Albert Dreyfus）上尉就是将法国机密泄露给德国人的间谍。他们需要一锤定音：证明清单上就是他的字迹。其实亨利和德克拉姆并没有别的证据，但是贝迪永不知道这回事，他带着反犹太主义的额外自信，带着认定犹太人德雷福斯有罪的前提，高兴地开始工作。

贝迪永为证明是德雷福斯写了清单，一刻不停地在思索着清单。他要研究清单和有德雷福斯字迹的样本，这两样东西存在相似性，但也有不同之处。这些不同之处足以让真正的专家推断出这两封信不是同一人所写。而贝迪永却“技高一筹”。他只要想象德雷福斯实施了如何高明的骗术，便可以通过这种对犯罪心理的品鉴来证明德雷福斯作为叛国者的罪行。

贝迪永纳闷：什么样的间谍会亲手写这种见不得人的信息？（其实真正的间谍会这样做，只是无人知晓而已。）贝迪永想到，德雷福斯既是间谍，还是犹太人，而犹太人那么精明，怎么会犯这么荒唐的错误。其字迹一定有所伪装。这就解释了为何德雷福斯正常的字迹与清单的字迹有所不同。

但现在贝迪永又遇到了另一个问题：如何解释相似之处呢？他这么

精明，为何不干脆使用一种完全不同的字迹呢？在探究这一问题时，他想到了其代表作，也是其体系的基石：自我伪造理论。

想象一下，精明的间谍会怎么做，贝迪永认识到，只有当可能涉案的文件在不会涉案的地点被找到时，变换字迹才有用。之后，德雷福斯就能利用这种差异声称他不是写清单的人。不过，若这封信在德雷福斯的身上或办公室里被找到，他就不能简单地声称不是他的了，作为骗人的高手，他会说自己被陷害了，是有人栽赃。但若有人试图栽赃德雷福斯，很显然会认真仿造其字迹。因此，德雷福斯开始模仿自己的字迹，即自我伪造。该事件发生时，芝加哥律师弗兰克·布莱尔（Frank Blair）撰文，嘲讽性地总结了贝迪永的推理：

> 简而言之，尽管贝迪永承认德雷福斯本人的字迹和清单的字迹存在不同之处，他却认为，这是德雷福斯上尉精心为之，以洗脱自身嫌疑；尽管在适当的情况下，字迹的绝对相似性反而使他能够声称这是对他本人字迹的模仿，进而断言是别人写的。[注]

若德雷福斯本人的字迹与清单的字迹相像，则有罪；若不像，也有罪。虽然推理存在明显瑕疵，但这并没有阻碍贝迪永，他在结论中声明："我得到了一组观察结果，还做了一些注解，详尽地容纳了所有事实，堪称完美，结论的可靠性毋庸置疑。"[注]因为有了贝迪永提供的"证据"，德雷福斯被捕了，接受了审讯，最终被革职并判处终身驱逐出境，被困在一个小岛的小囚室里，与法国相距甚远。

一年后，乔治斯·皮卡特（Georges Picquart）中校接管陆军情报机构。翻阅与德雷福斯的案子相关的档案，皮卡特找到了确凿的证据，证明真正的罪犯其实是军官斐迪南·华新·埃斯特哈奇（Ferdinand Walsin Esterházy）。他的字迹与清单的字迹完美契合。若贝迪永看到这一证据，也不得不承认埃斯特哈齐的字迹与清单的字迹有惊人的相似。但是

贝迪永从来不愁找不到理由支撑自己的观点，他声称，“犹太人一年来一直在训练人来模仿该字迹”，大概是想找一个替罪羊。[注]不仅是贝迪永，军队的高官也同样无动于衷。因为他们已经找到了罪犯，若罪犯另有其人的话，他们又得重新写报告，对他们而言只是添堵而已。基于皮卡特的发现，埃斯特哈齐被传唤受审。但在上将的庇护下，这个嫌疑犯被证清白，德雷福斯虽无辜但还是维持原判，结果是皮卡特遭到贬黜，德雷福斯仍被关在岛上。

这起案件引起了社会及政治上的大动荡，发起人是让·饶勒斯（Jean Jaurès）、利昂·布卢姆（Leon Blum）、乔治斯·克雷蒙梭（Georges Clemenceau）。最值得一提的是，小说家艾米尔·左拉（Emile Zola）撰写了题为《我控诉》（*J'accuse*）的公开信，以此给军队施压，让其重审案件。原判决于1899年宣布无效，同时，在雷恩（Rennes）开庭复审。贝迪永还是专家证人之一。但这次，任务的难度增加了，他必须要证明清单的作者不是草率的间谍埃斯特哈齐，不是他不加伪饰写下的便条，而是德雷福斯邪恶用心驱使下的精心设计之作。贝迪永能胜任这项任务。其证词密密麻麻地写满了50多页纸，还附加一些照片和证物，真是堪称其体系中的典范。

贝迪永考究清单中的每个词，权衡每个字母，拍下每个笔触，发现处处有迹可循。清单的第13行与第30行重叠，其中有3个字母是匹配的。[注]“利益”（intérêt）一词被取出来重写，两个版本首尾相接时，长12.5毫米，这正是军队地图的单位尺寸。[注]更证据确凿的是，这一单位尺寸的标准细分是1.25毫米，而这个词横看竖看都符合：以1.25毫米为计量单位，“t的长度横着是3；重音符的长度是1；长音符的宽度是1.5，最后一个t的高度是4，等等”。[注]这是巧合吗？不可能。清单一定是某个能工巧匠的杰作，用几个模板和军用直尺就创造了现代最复杂的伪造品之一。

这种考量毋庸置疑。在宣读了长达十个小时的法庭证词后，贝迪永

给出了有力的推论："一系列的观察与一致性形成了我毋庸置疑的证据，确凿的理论和实物证据都使我的断言更有说服力。像1894年时一样，我带着源自笃定信念的责任感，真心真意地认定清单是被告所为，并对上帝发誓，所言属实。我说完了。"[注]

我们很难知道全体审判人员对贝迪永这番证据的信任程度。不管怎样，法官还是再次宣判德雷福斯犯了叛国罪，尽管有"减轻处罚的情节"。这个荒谬的裁决反映了维持原判比厘清案件内在的是非曲直更重要，或者说根本就缺乏这种是非曲直。很明显，德雷福斯是无辜的，任何人都看得出来。与其等待一场可能永远不会发生的新审判，德雷福斯同意以接受有罪判决为代价，之后于1899年9月19日接受卢贝总统（President Loubet）的赦免。但他要最终官复原职、恢复之前的军衔，还得再等七年之久。

我们身上都有一个贝迪永

用推理来捍卫已形成的信念，贝迪永在这方面提供了一段引人入胜的故事。他似乎对自己的论据深信不疑。三位专家受指派对其体系进行评估，发现该系统"混乱不堪"且"完全没有科学价值"，令他们感到费解的是，"如此荒谬的逻辑究竟是如何形成那么大篇幅的阐述的"。[注]不过，他们也得出结论，"正是贝迪永揭开其体系奥秘时持有的天真憨厚之心，促使人们相信他确实是诚实无欺的"。[注]

贝迪永很容易让人当成疯子，与其同时代的许多人也确实这么认为。但是那样想就忽视了他在其他方面的辉煌成就，抹灭了其发迹于寒微、才思敏捷地擒拿罪犯的经历，这些，一个疯子不太可能做到。为了避免自命不凡的观点，我们还在实验室里再现了贝迪永思维模式的每个环节，展现了怎样的推理会把人引向不归路。幸好，实验只是在较小范围内重现了贝迪永的心理活动过程。这些实验明明白白地指出，推理才

是罪魁祸首。

当贝迪永提到自己考虑到了所有的事实，而且详尽程度简直完美时，明显过分自信了。根据唯理智理论，推理应该要使我们怀疑自己的信念，特别是当支撑自己的信念的依据站不住脚时，就像贝迪永遇到的那样。为何推理反而会导致过分自信呢？

亚瑟·科瑞特（Asher Koriat）30多年前就给出了答案。[注]在其某一项实验中，他让参与者回答常识性问题，如“科西嘉岛是法国的领地还是意大利的？”，由此详细说明这些参与者的自信程度。事实证明，他们都过分自信。就是说，当他们认为自己通常能答对80%时，其实只能答对60%。实验人员对某些参与者没有给出特别指示，而对另一些参与者则不同，他们让这些人给出支撑其答案的理由。这并不会影响自信，因为参与者一直都在那样做，即堆积支撑其起初直觉的理由。这就是他们过分自信的原因。而且，他们不像完全想不出理由来说明为何自己会犯错，若实验人员要求他们给出理由时，他们会应允，不会去推托，但给出理由时明显没有那么自信了。鉴于有自我中心偏见，他们没能自然而然地进行这种更客观的推理。贝迪永的第一份陈述并非完全确定，而从其15年后在雷恩的复审中所做的证词可以看出，他不仅更加自信，其信念也更加极端化了。在他眼中，德雷福斯变得愈加狡猾，且正筹划更大的阴谋，其证词更犀利，抨击力度大胜从前。15年之后，理性就将贝迪永推到了这种极端的地步，但其实，要察觉到极化的根源，所需时间要短得多。若参与者想某人想上几分钟，如果最初就喜欢这个人，他们最后会更喜欢他；但若他们最初不喜欢这个人，最后则会更讨厌这个人。[注]这段短暂的时间间隔让他们慢慢累积支撑并强化最初印象的理由。

起初，贝迪永证明德雷福斯有罪便是靠字迹间的相似性，这是他拿出的唯一证据。皮卡特后来给贝迪永看了埃斯特哈齐的字迹，这回才是完美契合，贝迪永本该立即改变看法的。但是当其最初印象遭遇皮卡特

的干涉时，贝迪永用理由搭成了一个不易变形的脚手架。这个脚手架支撑着他最初的主张，而不顾反对其主张的明确证据，这一过程被称为“信念固着”。

李·罗斯实验小组于1975年对此进行了演示，这是最早进行的实验。注实验人员让参与者辨别自杀便条的真假，接着会告知其表现好坏。紧接着又给他们几分钟时间思考自身表现。在这段时间内，他们想出了许多理由，来解释为何实验人员给出的反馈意见有道理，如他们总是很敏感、他们认识抑郁的人等。最后，实验人员告诉他们，其实反馈意见完全是个骗局，与其实际表现没有一点儿关系。但是这一切说得太晚了。参与者已找出许多理由来支撑自己的信念，所以即便失去了起初的动力，其信念仍能自行站住脚。有些参与者收到了积极反馈，他们认为，与收到消极反馈的参与者相比，他们做得更好，即便之后被告知反馈意见只是伪造的，他们仍如此认为。

拙劣推理并非个案

贝迪永可悲的例子说明：“若始于错误，还不知过错，逻辑学家就会变成疯子。”注他一直在堆砌理由，用它们来支撑其最初的信念，即认为德雷福斯有罪，他明显缺乏自我批评，所以最终才得到不切实际的荒谬的结论。贝迪永的想象力似乎无穷无尽，但毕竟只是他一个人。想象一下，当人们群策群力，都来找寻理由支撑自己的信念时，会产生怎样的结果呢？

阴谋论通常都是从某些微不足道、将信将疑的事实开始发酵的，譬如，月球上没有风，为何照片中插在月球上的美国国旗在飘动？世贸中心大楼真如人们所说，在“9·11”事件中被炸得粉碎吗？这些问题提出后，有些人上了网，他们找到宣传册，找到志趣相投的人。很快，这种疑惑不断升级，人们各执一词，尖锐的问题演变成完完全全的瞎猜。他

们认为，政府官员一定在撒谎。美国宇航局、美国中央情报局、美国联邦调查局、美国国家安全局这些政府机构一定牵涉其中。他们甚至还认为，该阴谋必定是全球性的，可能受到联合国的推动，也可能是彼尔德伯格集团（Bilderberg Group）一手策划的，甚至极有可能是“国际犹太人集团”所为，这个集团传承了贝迪永时期的反犹太主义阴谋论。注

尽管阴谋论甚嚣尘上是成百上千人推波助澜的结果，但是人数少不代表就不会被理性引入歧途。在20世纪60年代，耶鲁心理学家欧文·贾尼斯（Irving Janis）开始对此进行研究，目的在于搞清楚这些人何时会做出不明智的选择及其原因。1961年美国政府袭击古巴未遂，变成声名狼藉的猪湾侵袭，贾尼斯对此进行了认真分析，分析是什么导致这一致命的决定。他认为，群体思维才是罪魁祸首，群体成员没有批判意识，没能批判性地看待彼此的建议，也没有考虑备选项。注

心理学家在实验中不仅看到了缺乏异议所造成的问题，还收集到群体极化的证据。把一群人放到一起，让他们谈论所赞同之事，结果会增强某些人的信念。有种族偏见的人偏见会更强，主张人人平等的人平等意识会更强。注主战的鹰派（Hawks）将进一步增加对军队的支持，而主和的鸽派（Doves）则将进一步减少对军队的支持。注若你赞同某人所言，你不会认真考究他的论据，毕竟，你已经认同他的结论了，所以为何还要多此一举呢？若想法一致的人在一起辩论，他们会彼此给出支撑已持有之信念的新理由。就像孤独的推理者，志同道合的人也会受到信念极化、过分自信和信念固着的影响。注

进化论不在乎我们感觉有多好

推理可以让人变得稀奇古怪，就如贝迪永不断发展完善其体系，阴谋论者认为蜥蜴人将统治地球一样。我们不屑地将持有这种信念的人视

为疯子或笨蛋，这再自然不过了，或者，用一种更礼貌的口吻——他们的认知受到了局限。

这一解释没有得到事实的支撑。因为贝迪永既不是疯子也不是笨蛋。要知道，极少数阴谋理论家会罹患精神病或有认知障碍。[注]此外，尽管我们不能怀疑有些人存在任何智力缺陷，他们似乎遭遇着同样的困境。莱纳斯·鲍林一开始只是把维生素C当作治疗普通感冒的药物，到最后才将之誉为万灵药。越南战争不断发酵，背后黑手都是些“智力超群、受过良好教育、消息灵通、经验丰富、爱国且得力”的领导人。[注]

这就对唯理智理论提出了疑问。唯理智理论认为，推理应让人改正错误直觉，可实际上推理不仅没有做到，还让人不论正确与否，都更坚定不移，让人没有合理的理由就敢坚持自己的信念。众多历史案例证明，这些不是经巧妙的实验被放大的小怪癖，而是会带来悲剧后果的真实现象。

有时，心理学家会利用认知性解释与动机性解释之间的区别。做错事或信念有误的人要么是因为错误认知，要么是因为受到了刺激才会误入歧途的。既然错误认知解释不了推理所产生的惊人结果，明智之举就是像基瓦·孔达在抵御动机性推理泛滥成灾时那样，提供动机描述。她认为，理性之所以把人引入歧途，是因为有“意愿、渴望或偏爱”的驱动，让人去持有注定要有的信念，不论其正确与否。[注]理由林林总总地摆在那儿，由不得人们不去相信，但大多数人竟选择自我麻痹，因为这让他们感觉良好。受到渴望的驱动，“与令人不快的观点和记忆相比，人们更愿意相信让人舒坦的观点和记忆”。[注]

就其极端而言，感觉良好的解释可以简单地认为，人们会接受任何使之感到愉悦的信念，这是“主观臆断”（wishful thinking）的极端形式。人们普遍使用“主观臆断”来描述某人持有某观点并不是因为有证据证明其正确，而是因为自我的渴望。而且，在普通的对话中，人们通常

不只将“这是主观臆断”当作一种描述，还当作一种解释，譬如“为什么约翰认为自己很受欢迎？这只不过是他的主观臆断，他希望自己如此”。

从心理学视角来看，甚至更多的是从进化视角来看，主观臆断的想法有待解释，而不是作为一种解释。信念要指导行动，就应该先让我们对世界有个了解。若世界并非想象的那样，我们最好及时意识到，这种错位正好能让我们有所作为。有观点认为，事情之所以按人们希望的样子呈现，只是因为人们这样期望了，这种想法违背了信念的主要功能。

因此，孔达及其他动机推理专家的举动就没什么可奇怪的了，他们指出，人们不会因为自认为恰当就去接受某观点。他们会去找寻理由，并确保能为之辩护，若没能找到正当理由，便会不惜摒弃其在意的观点。譬如，人们基本上都更愿意相信自己比一般人优秀，认为自己更聪明、更擅长社交、更明智等。不过，他们不会简单地认为自己的一切都是最好的，就因为这让他们欢喜若狂，而是有选择地自我拔高，当其在某种程度上能站得住脚时，才对自己给出过高的评价。

譬如，人们通常会觉得自己比一般人更聪明，这个观点比较容易辩护：他们可能擅长数学，善于营生，富有教养，长于社交，等等。相对而言，有的观点则不容易辩护，如不存在两种“守时”的方式。既然人们想不出自己如何才能比“一般守时”更守时，就只有摒弃该观点，对其他同样难以辩护的观点亦是如此。注

即便推理不是一厢情愿的想法，让我们感觉良好仍可以作为其功能。该假设首先指出推理经常会产生相反的影响。贝迪永可能会目空一切地自信，会迷失在这种喜悦当中，也可能因为德雷福斯诡计多端，因为出现的每个新证据而苦恼。记者乔纳森·凯（Jonathan Kay）在写作《真相持有者》（*Among the Truthers*）时，采访了许多阴谋论者，即不接受“9·11”事件的官方说法的人。他发现，人们由于“突然暴露于强大

的、威胁世界的邪恶力量之下”，[注]起初是陷入极度痛苦，而后才慢慢缓解。如果推理的任务是让人感觉良好，那么它在这方面失败得很彻底。我们很容易找到其他例子，譬如，善妒的丈夫坚信妻子在骗他，悲观主义者不断找寻理由说明为何人类注定要自我毁灭，疑病症患者在自己身上找寻其他疾病的症状。[注]

令人感觉良好的解释不仅不符合事实，从进化论角度看也没有意义。它混淆了解释的两种层次，即最近似解释与最终解释。[注]最近似解释旨在查明某行为的心理学或神经病学原因。譬如，迈克尔口渴了，就喝了点水，他从喝水中获得的乐趣就是对喝水这一行为的最近似解释，就是说，他喝水是希望能感觉舒服。

相对而言，最终解释从进化论角度来解决问题。归根结底，感觉舒服只是达到目的的手段而已。从进化论角度来看，享乐状态（诸如愉快、痛苦、幸福、绝望等）能刺激动物做出某些举动，而这些举动对其生存与繁殖都至关重要。喝完水后我们感觉身心愉悦，因为水是生存必不可少的。若触碰正在燃烧的圆木，我们会感觉灼热，因此缩回了手，以避免长时间的损伤。我们喜欢与朋友一起玩，因为对人类繁衍生息、不断进化而言，伴侣和盟友发挥着十分重要的作用。同样，若遭到朋友的抛弃，我们会很绝望。若某个个体觉得饮水很痛苦，觉得手被烘烤很享受，又或者他讨厌被人喜欢，酷爱被人厌恶，这样的人要繁衍生息就会很困难。所以，不论推理是否让人感觉舒服，都不是为了这个目的。

推理中的适应性滞后

我们运用互动理论，可以解释理性造成的各种认知扭曲，即过分自信、信念极化和信念固着。在第十一章和第十二章中，我们指出了理由制造的两个主要特点。第一，它是有偏见的，人们会不顾一切找寻支撑其观点的理由；第二，它是懒惰的，人们不会仔细检查自己的理由。这

两个特点加在一起会阻碍孤独的推理者。他在推理时，会找寻越来越多的论据来支撑自己的观点，而且认为大多数论据都是足够充分的。这些理由虽增加了其信心，但也把他引入极端之境。

许多心理学家可能会认同这一诊断结论。不过，给出这个解释应该只是第一步，人们很快就会有疑问：推理到底为何要如此表现呢？若手工艺品不尽如人意，可能是因为破损了，也可能是因为其操作条件异常。若你的钢笔倒过来写不出字了，你的小轿车油箱空了不能启动了，不是因为它们发生故障了，而是因为其操作条件变了，与设计不相符了。生物装置也要在其所适应的正常条件下运作。[注]近地面的大气压形成了人类肺部的正常运作条件。我们的肺部能在这种条件下正常运作，但若条件异常，运作就会有障碍或根本不能运作，如在高海拔下、水下、在装满氦气的箱子里就不能运作。

在互动理论中，推理正常运作的条件是社会环境，具体地说，是对话环境。脱离这一环境，推理者的利益就不能在推理中得到保障。认知扭曲和失策之举都是有可能发生的。这并不是因为推理失效了，而是因为其运作条件异常。同样，在地下停车场钠灯的照射下，我们会看到物体呈现新的色彩，这并不是因为我们的色觉失灵，而是因为运作条件异常。我们的色觉被人造灯左右了，毕竟人类演变过程中主要还是利用太阳光，而且已适应了太阳光。

当然，若单纯认为推理现在的运行环境通常都是异常的，还不够全面。如果炸弹在轰炸机内爆炸，而不是在击中目标时爆炸，主管工程师就算指出爆炸恰好发挥了预计的力量，也不会因此受到嘉奖。炸弹爆炸的时间与方式一样重要。同样，认知机制发挥作用的时间与发挥作用的方式至少应该同样重要。

和对话方发生观念冲突后，一般会触发推理。我们会试图构造说服他人或至少维护自身立场的论据。即便没有真正的对话方，推理也可能

会被触发，因为要预测可能产生的分歧。预测有时会十分具体，如定好会议安排、努力解决某分歧或就双方观点进行讨论。平时，我们可能只是预想会与政敌不期而遇，会在心里准备并演练后来会急于使用的论据。有些时候，我们甚至会重演发生过的辩论，并认为我们早就该用这些论据。

打个比方，萨沙（Sasha）准备得到母亲的允许，去参加瓦内萨（Vanessa）家举办的通宵派对。他一直在演练自己的论据：他在学校表现很好；他下周的家庭作业也做完了；聚会只是件小事，他不会做出格的事，应该不会让母亲担心。越这样想，他就越相信自己的请求十分合理，而且母亲理应会同意。

接着会出现好几种情况。萨沙可能会说服母亲同意，毕竟也没什么好反对的。他母亲也可能会说服他，认为这终究不好，因为她从其他父母那里听到，聚会上可能会有比他们大的孩子擅入，这些人会带上啤酒，甚至还可能会带毒品。另外，他似乎忘记了下周还有一场考试，对于这场考试他至今都没有准备好呢。听着母亲给出的论据，萨沙可能想要反驳。这样反反复复，终究会有一个人说服另一个人，或者说，二人至少会拿出理由来辩护自身立场。

相对而言，如果萨沙的母亲只是说不行，而没有关注萨沙给出的论据，又或者萨沙不敢去问母亲，他会自认为自己的论据令人信服，他会觉得父母不公、不理解他，觉得自己是个受害者。因此，只要争辩真正会发生，在预期有争辩时进行推理是可取的。

单个推理本身没有问题，有问题的是其单一性。然而，推理一定会时不时逗留在人的大脑中，因为人们不能完全预测到何时会被叫去捍卫自身观点。正因为人们会被预料之外的理由诉求惊到，才要早做准备，尽管立场的冲突可能永远不会发生，也要如此。由于两种类型的错误之间成本的差异，后者很可能更加普遍。若持有他人反对的观点或行为，却没有早做准备，预想辩护之词，一旦被抓包，与预先演练辩护之词相

比，情况会更糟糕，尽管预先演练的辩护最终可能用不到，但这种现象更普遍。

现代环境扭曲了我们预测分歧的能力。这只是许多情况中的一种，许多情况下，环境变化太快了，自然选择很难跟上步伐。譬如，咖啡、香烟和酒这些刺激神经类物质都是在现代环境下产生的，且成为普及之物。其中香烟明显对吸食者的身体有害。然而，我们还没有进化出一种与生俱来的天性，让我们不去使用这些物质，就像我们有与生俱来的天性，让我们不去触碰有毒的食物。可以说，这些物质本该在我们的进化过程中越来越罕见，但因为它们最近变得随处可见，所以我们不得不适应这一变化。

环境变化是否让我们预测分歧的能力变得失衡了，就像我们让自身对精神刺激物质的反应变得危险一样？与远古环境下的生活相比，富饶的现代社会下的生活在很多方面都是不同的，其中一些不同之处必定会影响我们的推理方式。例如，在印刷术的发明和现代媒体的诞生之前，人们得知团体内某人观点与其不同，通常是通过与那人的互动。我们能找出观念的差异并试图解决，通常是因为反复地交流论据，而且我们可以预想这些论据并在脑海中演练。现在，我们身边充斥着主笔、新闻节目主持人、博主等人的观点，我们也许永远都不会与这些人见面。当然，别人也希望我们能就许多不同的话题发表看法，政治、音乐、食物都行，还要当其受到质疑时予以辩护，尽管那些辩论可能永远也不会发生，但我们还是要拿出理由，以备不时之需。

然而，这只触及了问题的表面。还有更多急剧的变化会影响理性的运作。有些人住在大城市，他们每天遇到的陌生人都多于其祖先一辈子遇到的陌生人。其中许多陌生人都有不同的文化背景。很容易判断，这种异常的融合会带来诸多产生分歧的可能性，而要想出合适的正当理由变得比从前复杂得多。

一些认知机制在现代世界已经完全改变了意图，所以与其远古形式

只存在少许相似之处，它们见证了读写能力对我们的改变，让我们能辨识任意的简单图形。[注]虽然我们认为理性没有经历这种剧烈改变，但是环境变化无疑已经影响了触发理性的时间、理性发挥作用的方式，以及理性要实现的目标。理性现在的用途多种多样，与其进化来的功能有所不同，可以用来在正式辩论中展现聪明才智，也可以用来揭示物理规律。令人遗憾的是，理性的一些新用途，如为可能永远不会发生的辩论做准备，竟然对推理者有潜在的危害。正如凯恩斯（Keynes）所言："如果一个人独自思考太久，很难想象他一时会轻信怎样愚蠢的事。"[注]

---

1. Arthur Conan Doyle,The Hound of the Baskervilles, Project Gutenberg, 网址：http://www.gutenberg.org/files/3070/3070-h/3070-h.htm。
2. "Bertillonnage", 维基百科，网址：http://fr.wikipedia.org/wiki/Bertillonnage。
3. Blair 1901, p.395.
4. Reinach 1903, Kindle locations 2911-2924.
5. Whyte 2008, p.78.
6. Anonymous 1900, p.330.
7. 同Anonymous 1900, p.344。
8. Anonymous 1900, p.362。
9. Anonymous 1900, p.369。
10. Anonymous 1904.
11. Anonymous 1904.
12. Koriat, Lichtenstein & Fischhoff 1980。要注意到过分自信是个较为复杂的现象，推理并非其唯一的原因。
13. Sadler & Tesser 1973;欲了解相关评论，可参见：Tesser 1978。
14. Ross, Lepper & Hubbard 1975。欲了解推理的作用，可参见：Anderson, Lepper & Ross 1980; Anderson, New & Speer 1985。
15. Keynes 1989, p.252.
16. 可参见：Kay 2011。

17. Janis 1982.
18. Myers & Bishop 1970.
19. Myers & Bach 1974.
20. Sunstein 2002.
21. 例如，Kay 2011。
22. Halberstam 2001, p.xi的序言，作者约翰·麦凯恩（John McCain）。
23. Kunda 1990, p.480.
24. Nickerson 1998, p.197.
25. Dunning, Meyerowitz & Holzberg 1989.
26. Kay 2011, p.217.
27. 可参见：Scott-Kakures 2001。
28. Scott-Phillips, Dickins & West 2011; Tinbergen 1963.
29. Millikan 1987, p.34.
30. Dehaene 2009.
31. Keynes 1936, p.vii.

第十四章
普适性理由

我们从第十三章中可以看到，孤独地运用理性会把人引入歧途，这是因为他们一开始有强烈的直觉，如认定德雷福斯有罪、认为这是解决问题的正确做法等。自我中心偏见与宽松的评价标准让我们堆砌了许多肤浅的理由来支撑自己最初的直觉，而不论其对错。不过，我们一开始通常并没有形成强有力的直觉。对某些话题，我们只有微弱的直觉或根本没有直觉，如见到超市的走廊堆满清洁剂或厕纸时，我们的感觉很普遍。我们有时也有强有力但冲突的直觉，如是经济学还是生物学呢？艾伦（Allan）还是彼得（Peter）呢？在家陪小孩还是回去上班呢？

这些直觉应该是让个人主义理论大放异彩的有利条件。理性有绝佳的机会扮演公正的仲裁者。若推理者没有明显的偏见，就能提前避免自我中心偏见，理性就能对推理者的选择起到指导作用，或许还能起到积极作用。也许就是在这些情况下，哲学家才想出了有关理性效能的观念。他们觉得自己有责任认真查看直觉微弱或相冲突时的例子。而且若冲突不够充分，哲学家也很擅长引起冲突：你确定有其他人存在吗？再想一下！有一个哲学理论叫唯我论，就是认为其他人不存在或至少你永远不能确定他们存在。理性符合经典理论的预期，对这一点，没有直觉、有微弱直觉或有相冲突直觉的情况都能为之提供完美范例，经典理论期望在两种不同直觉间维持现状，认为只有那样做才能让理性发挥作用。让我们看看在这种情况下理性有哪些作用，先从拉斯·霍尔、彼得·约翰逊及托马斯·斯特兰德伯格（Thomas Strandberg）开展的巧妙实验开始探讨。

你走在大街上，一个年轻人拿着剪贴板向你走来，问你是否愿意参加一项小调查。这一次你同意了。他递给你的剪贴板上附有两页纸，上面写的是有关政治、道德和社会问题的陈述，如“若某举动可能会伤害无辜的人，要是做了的话会受到道德的谴责”。你必须表明对每段陈述的感觉，从“完全不同意”到“完全同意”选一个。你填好后就把剪贴板递了回去。可是你的任务并没有完成，这个年轻人又把剪贴板递了回来，让你解释其中某些评价。你很乐意这样做，毕竟作为一名见多识广、体贴周到还有主见的市民，你还是很自豪的。

你没有意识到的是，在几秒内，这个年轻人（他实际上是个实验人员），就在他拿着剪贴板时，已经通过简单的把戏把页面上某些陈述变成了恰好有相反意义的陈述。例如，关于伤害无辜的人的这段陈述现在变成了“若某举动可能会伤害无辜的人，要是做了的话道德上是允许的”（用“允许的”代替了“应受谴责的”）。若某些陈述你一扫而过，只是认真看了看选项，对于这些陈述，这张纸现在则会显示，你持有与先前恰好相反的观点。若你先前表明强烈支持第一种陈述，这张纸现在则显示你强烈支持第二种陈述，也就是相反的情形。

只有半数以下的参与者注意到这些新选项的异样。大部分人继续为与先前立场相反的立场辩护，特别是当其立场不太坚定、很容易辩护时。(注)

我们能无限地制造理由，但正如霍尔及其同事展现的那样，这只是为了支撑我们相信之事，或者说我们认为自己相信的事。自从理查德·尼斯贝特和蒂莫西·威尔逊于20世纪70年代开展开创性实验（该实验我们在第七章中已经援引过了）以来，我们的这一能力已变成社会心理学的主题。在其中一项实验中，尼斯贝特和威尔逊站在商场外面，假装卖长筒袜。(注)一些路人在他们的摊位前停了下来，选了自己喜欢的，当问及其选择缘由时，这些路人也很乐意为自己的选择辩护：“这种袜子看起来更耐穿”“我更喜欢这种袜子的颜色”。但是心理学家明白，所有的

解释都是伪造的，因为他们恶作剧地摆放了严格意义上完全相同的袜子。就算所有的长筒袜都相同，也没有阻止人们表达其偏好，这肯定是因为这些袜子摆放的位置不同。譬如，许多参与者就表现出“最右偏见”（right-most bias）。

在这类实验中，参与者最初只有微弱的直觉。在第一项实验中，参与者没能察觉到陈述被替换，是因为他们先前只是平和地表达认可或反对，没有对问题形成强烈的直觉。在第二项实验中，长筒袜全都是一样的，所以不论参与者有何偏好，都是为维护自身立场杜撰出来的，因此不太能站得住脚。但是，理性并没有服从经典理论的安排。它并没有对情况进行客观的评估，因而不能指导推理者做出更合理的决定。它找寻理由只是为了支撑稍微强一点的直觉。人类是理性化机器。正如本杰明·富兰克林（Benjamin Franklin）所言，“讲理能为人带来许多便利，它让人能找到或制造理由来支撑自己的所有行动”。[注]

然而，在某些情况下，理性确实会影响人们的决定，不过，这一理性并不符合唯理智理论。

推理何时发挥作用

我们又找到威尔逊的一项实验，但这次不是长筒袜，而是海报，也没有花招，因为所有的海报都是不同的。这项实验比较直接。实验人员让其中一些参与者评价五张海报，仅此而已。让另一些参与者也评价这五张海报，但要给出解释。[注]让参与者运用推理会影响其选择，正如他们会因此给幽默的海报更高的评价。

其他许多研究也表明，诉诸理性会有所不同。有些实验让参与者为其决定辩护；[注]另一些则给参与者额外的时间来思考自己的选择；[注]也有一些实验把基于感觉的决定与基于推理的决定放在一起比较。[注]每

次，运用更多推理的人的表现都与推理较少或根本不推理的人的表现不同。按说纯粹的理性化机器不会影响决定。所以，到底发生了什么呢？理性在帮助人们做出更好的决定吗？

不完全是这样。在每项实验中，推理得越多，得出的决定就越糟糕。譬如，在威尔逊的实验中，实验人员让参与者把其评分最高的海报带回家。几周后，当实验人员问及参与者对海报的欣赏程度时，与依靠本真的直觉进行评价的参与者相比，要对自身偏好进行解释的参与者满意程度要低一些。

要理解为何即便自我中心偏见不是罪魁祸首，理性还是会打乱人们的决定，我们必须更认真地观测理性影响决定的方式。

伊塔马尔·西蒙森（Itamar Simonson）很早就围绕该话题开展了实验。[注]他一开始便拿两个啤酒品牌来打比方，而且这两个牌子深得大多数人的喜爱。第一个牌子是德涞宝啤酒（Beer Deluxe），是一款高档啤酒，对其品质评分是75分（满分是100分），半打18美元。第二个牌子是比罗斯（Beeros），品质评分是65分，但是较便宜，半打只要12美元。这让人很容易取舍。当参与者要在这两个牌子间做出选择时，他们想都没想，就拿了经常喝的那种。接着实验人员又介绍了第三个牌子——高级啤酒（Premium Beer），半打18美元，和德涞宝啤酒一样贵，但品质评分没有那么高，只有70分。鉴于高级啤酒明显比德涞宝啤酒差，应该不大会影响人们做选择。但其实是有影响的：一旦人们知道了高级啤酒这个牌子，会更倾向于选择德涞宝啤酒。

克里斯托弗·哈西（Christopher Hsee）的一项最新颖的实验就是围绕该领域开展的。他问参与者，若任务完成，有两个礼物供他们挑选，他们会选哪一个作为奖励。两个礼物都是巧克力，其中一个较小（0.5盎司）而且便宜（0.5美元），是心形的；而另一个则较大（2盎司）而且偏贵（2美元），是蟑螂形的。若参与者更多的是依靠感觉，则会分

成两派。但若经过推理，大多数参与者则会选择那个较大的、蟑螂形状的巧克力。注

黛博拉·汤普森（Debora Thompson）探究了功能蔓延（feature creep）现象，即增加无用特点就会让设备负载许多小配件，最终只会降低其实用性。她与同事迈克尔·诺顿（Michael Norton）共同表明，当人们感觉到必须要拿出支撑自己决定的理由时，会更倾向于选择充满特点的物件，比如有几十项功能的数字视频播放器，即便他们意识到使用起来会很不方便。注

所有实验结果共同的脉络是：在每个案例中，理性都会促使参与者做出更容易辩护的选择。“德淶宝啤酒品质更好，也不比高级啤酒贵，所以我会选择德淶宝啤酒。”“鉴于它是个礼物，干吗不挑更大更贵的呢，因为外形而不选是不理性的，无论怎样它又不是真的蟑螂。”“为什么要买功能较少的数字视频播放器呢？”

这一普遍现象就是知名的推理性选择：当人们的直觉较微弱或相冲突时，理性会促使他们做出决定，因为这样最容易找到理由来支撑，他们也可以为之做出最佳辩护。

为理由付出代价

虽然这些实验结果很难与唯理智理论达成一致，因为理性本应促使人们做出更好的决定，而不是更糟糕的决定，但是符合互动理论的预言。就算不用维护立场，理性也一直是社会性工具。它会针对不同的选择，筛选出各种理由，然后促使推理者选择最容易辩护的，而不管其好坏。

在许多情况下，推理似乎会让人做出更糟糕、不太理性的决定。引

入明显次要的选择（如高级啤酒）不应该会对两个优质选项间的选择有影响。对厌恶有所研究的心理学家会告诉你，不论那个蟑螂形状的巧克力有多大，人们吃它时是不会感到享受的。[注]充斥着无用特点的装置会造成焦虑，而不是享受。[注]就因为有更多种果酱可供挑选，就不去买某一种果酱，这种理由没有太大意义。

更让人震惊的是，人们只愿意为有理由支撑的决定买单。阿莫斯·特沃斯基和埃尔德·沙菲尔（Eldar Shafir）是第一批与伊塔马尔·西蒙森一起探索推理性选择的研究员，他们让第一组参与者设想下面的场景。

> 你刚参加了一个费力的资格考试。正值秋季学期末，你身心俱疲，而且你还不确定是否通过了考试。万一没过的话，几个月之后又要参加一次，考试时间在圣诞假期之后。你现在有机会买到极具吸引力的五天圣诞度假套餐，度假地点在夏威夷，关键是价格非常低。这个特惠活动明天截止，而考试成绩也是明天揭晓。
>
> 你会：
>
> 1. 购买度假套餐。
>
> 2. 不购买度假套餐。
>
> 3. 支付5美元的不予退还的服务费以保留购买度假套餐的权利，后天还能享受同样优惠的价格，也就是在你弄清楚到底通过考试没有之后。[注]

实验人员让第二组参与者设想已通过考试，让第三组参与者设想没通过考试。

大多数通过考试的参与者都决定购买度假套餐，因为他们觉得这是其通过考试应得的嘉奖。大多数没通过考试的参与者也决定购买度假套餐，因为他们觉得自己迫切需要休个假，好从考试失利中缓过来。

结合这两组实验结果来看，对于第一组仍不知考试结果的大多数参与者而言，购买度假套餐无论怎样都是理性的，不必浪费5美元来推迟做决定。也就是说，不论过没过，他们都会购买度假套餐。然而，该组参与者却选择支付服务费推迟两天特惠期，就为了知道考试结果。他们的问题在于，认为支持购买度假套餐的两个理由不能兼容，一个是“这是我通过考试应得的嘉奖”，另一个是“考试失利后我需要休个假”。所以，他们花钱来等结果，实际上是在买理由以做出无论怎样都会做出的决定。

社会理性

显而易见，当我们对决定进行评估时，应该关注其内容与实际目标之间的符合程度。我们应该买自己更喜欢的海报。购买电子设备时，我们也应该购买最符合需求的样式。然而，如果变得理性就是要努力做出最好的决定，要把所有事情都考虑进去（而不只是我们的实际目标），那么做出好的决定就变得更复杂了。

人们时常相互评估。与我们互动的人有能力、可靠吗？他们的判断合理吗？正如我们在第七章中主张的，许多这类评估都是根据理由好坏做出的：我们通过理由归因来理解他人的想法及行为，评估他人所给理由的合理性，进而评估其可靠性。我们对理由的依赖可能歪曲且夸大了其对想法及行动的作用，但这并不是说更好、更客观的理解就一应俱全了。毕竟心理学家本人还在努力形成这种理解，而且还对这种理解的形态持不同意见。尽管基于理性的理解存在许多缺点，但还是可以用来解决两个主要问题：一是在评估人们表现时提供依据，二是为人们表达、分享并讨论这些评估提供惯用说辞。

就像我们评估他人一样，他人也评估我们。对我们而言重要的是，他人对我们有一个好的看法，因为这会促使他们更愿意与我们合作，而

不太会做不利于我们的事。鉴于此，应该做出有效行动，不仅是为了实现我们的目标，也是为了维护良好的声誉。让我们如此做的理由不应该只是合理的理由，而应该是很容易就公认为合理的理由。

在某些情况下，我们拿出的个人最好的理由可能太难懂了，或可能与常识观念相悖，因此，反而会对我们的声誉有损。这种情况下，更有利的选择是做出更容易辩护的决定，它比最优决定可能要差那么一点，却是最合适的。我们可能得不到实际的好处，却提高了社会声誉，总体而言，好处会更多。

就获得声誉而言，理性会影响我们的决定。譬如，参与者之所以选择德涞宝啤酒，是因为这个决定最容易辩护，这款啤酒可能不是他们最满意的，却让他们获得了社会声誉，因为这样的决定最不可能招致批评。(注)有的消费者最终选择了有许多无用特点的东西，会被嘲讽为“技术行家”。(注)努力看起来理性，甚至不惜做出一些实际并不理性的举动，这可能是最理性的做法。

就我们在这章已检验的选择类型而言，人们的直觉通常都是微弱的。拥有这种微弱直觉通常是个可靠的迹象，因为这表明尚在讨论的抉择没那么重要。毕竟，当谈及处理我们的祖先生活的环境中最迫切的问题时，具体的认知机制很可能已经进化了，可以为我们提供强有力的直觉。所以，一般情况下，当我们只有微弱的直觉时，以辩护某个特定抉择的难易度为指导，就是简单且合理的启发法。

然而，我们生活的环境在过去的几千年、几百年甚至是几十年里变化很大，所以微弱的直觉不再是对抉择真正重要的可靠指示。譬如，对大多数人而言，买车是个重要抉择。因为他们要花许多钱来买车、给车做保养，要花许多时间来开车，还要把生命与车的安全性能绑在一起。虽然我们已有认知机制来帮忙选择安全的食物和可靠的朋友，但是还没有机制进化来帮我们选车。因此，直觉只能提供微弱和有限的指导。难

道这意味着在这些新奇的进化情境中，找寻容易辩护的抉择，如选择一款受欢迎、口碑好的车，是不合理的冒险吗？不见得。在大多数情况下，在他人眼中最容易辩护的决定、最可能提高我们声誉的决定，也是实现我们目标最好的决定。

在给定的群体中，公认合理的理由，若客观上也是合理的，关心声誉的人还是会形成正确的观点、做出有效的决定。但是情况并不总是这样，远非如此。几百年以来，衣冠楚楚的内科医生做出夺走病人生命的决定时，竟觉得理由正当、问心无愧。对盖伦（Galen）的体液学说等生理学机理的错误理解，导致像“给病人放血来恢复体液平衡”这样容易辩解的决定与“某些病人放血后病情恶化”这样的事实错位了，这样的事实必定使其中某些医生踟蹰不前。但是，如果他们急切地想要维护声誉的话，最好还是继续给病人放血，毕竟也没有其他的补救措施。相对而言，如今的医生都是依靠极大改善的、基于证据的医疗知识，所做决定也是得到医疗界认可的，这样既能维护其声誉，又不会在其错误指导下危害父母的身体健康。

辩护与论证何时分化

本章的信息可能看起来比较单薄。理性帮助我们提高社会地位，而不是引导我们做出本质上更好的决定。而且即便它引导我们做出了更好的决定，主要还是因为我们所处群体碰巧赞成对这一问题的所谓正确决策。不过，这不能代表全部。就理由而言，辩护的确涉及尊重常识或专家意见。不过，暗中为这一尊重之举辩护的是“群体或专家更擅长于拿出合理理由”这一猜测。可是，社会公认的“合理理由”所提供的这种懒惰的辩护，与更好地理解、评估这些理由并自学一些专业知识的个人努力之间存在冲突的可能性。

举个例子，19世纪初，巴黎最德高望重的医学权威是约瑟夫·维克

多·布鲁塞斯（Joseph Victor Broussais）医生。他坚持认为，发烧全是由炎症引起的，并认为都应通过放血来治疗。其晚辈皮埃尔-查尔斯-亚历山大·路易斯（Pierre-Charles-Alexandre Louis）医生并没有真的怀疑放血的功效，只是想给出确切的评估而已。为此，他比较了两组因肺炎已被放血的病患，发现与预期相反，较之放血较晚的患者，在患病早期就被放血的患者死亡人数要多得多，这一现象表明，放血不仅不能治愈病患，反而会使其身体状况恶化。路易斯现在有了令人信服的证据及论据，即便他总体而言不反对放血，但至少对布鲁塞斯建议的系统性放血还是反对的。路易斯在实证医疗中的开创性研究，在19世纪逐步摒弃放血这一主要医疗步骤的过程中发挥了关键性作用。路易斯冒着瞬间身败名裂的风险，批评过分扩张的放血实践，但恰恰因为他有合理论据的支撑，才最终获胜，还提升了自身声誉。(注)

如果我们能这么想就好了：当拥有他人眼中合理理由的目标与拥有可论证的合理理由的目标之间发生冲突时，论证的说服力力压即时辩护的安逸；最好的理由最终会胜出。好吧，事情变得稍微有点复杂了。

设想一下下面的场景。

> 你有一张篮球赛的票，这场篮球赛在距离你家60英里（约96.6千米）的城市举行。比赛的这天有暴风雪，道路非常不好走。以下哪种情况让你去看比赛的可能性更大：
>
> 1. 你买票花了35美元。
> 2. 票是免费的。
> 3. 二者皆有可能。(注)

大多数人的回答是，若他们是花钱买的票，而不是免费得到的，则更可能会冒着暴风雪去看比赛。心理学家哈尔·阿克斯（Hal Arkes）和彼得·艾顿（Peter Ayton）认为，做出这一决定的理由是因为“浪费是可

耻的”。注这样，大多数人就能更好地理解为何票花了钱会让人不惜冒着暴风雪去看比赛了；他们甚至可能对那些买了票但没去的人不敢苟同。

一方面，经济学家把该场景中的票钱称为“沉没成本”。钱已花，不能收回了，这和“沉没”没两样。决定是关于未来之事的，可以更改，不是关于过往之事的，过往之事就不能更改了。那么，唯一且真的重要的问题是：你怎样做会更好，是冒着暴风雪去看比赛，还是做些其他事呢？若做其他事会更好，做令人不快且可能有危险的事（去看比赛）则只会更糟糕。认可这一论点的人似乎应该回答，不论票花钱没有，都不会影响他们的决定。

另一方面，反对所谓沉没成本谬论的论点也很明确。该论点甚至还得到了许多哲学家及经济学家的支持。不过，如果你对反对沉没成本谬论的论点深信不疑，即便花钱买了票，还是决定待在家里，可能会受到他人的指责，那些人不知道该论点的存在，你可能也没机会解释。所以，人们可能很讽刺地认为你做了一个不理性的决定，而实际上有力的理由支撑着你的决定。也就是说，若你在乎他人的看法，则会做出社会认可的决定。注（注意，如果你在学经济学，并基于沉没成本做决定，那么你的家人是不会认为你判断不当的，但你学经济学的同学可能最终会认为你无能。）

可是，为什么沉没成本谬论竟会变得普遍，甚至还不被当作是荒谬的，而是被当作要做的正确之事呢？这里有一个猜测的答案：人们在朋友、搭档或合作伙伴身上找寻的其中一个品质便是可靠性。在执行任何决定时，在继续实施任何行动方案时，即便调低了预期收益，一定程度的执拗也可以向他人证明其可靠性。有些人持之以恒，即便从其个人立场来看，这可能不是理性的最佳选择，但这样做增加了其可靠性，提升了其声誉。那么，至少在某些情况下，理性之举可以是有意承认沉没成本谬论来向他人展示自己的可靠性，还可以是与不承认该谬论的人相

比，对承认该谬论的人做出更高的评价。

关注理性的互动功能不仅让理性更有意义，还展示了其局限性。辩护性和论证性理由是人类互动的基本工具，但当其分化时，哪种理由更胜一筹，可能不仅取决于理由的质量，还取决于社会利益，特别是声誉。理性模块不能假装占据经典理论赋予理性的主导地位。

---

1. Hall, Johansson & Strandberg 2012.
2. Nisbett & Wilson 1977.
3. Franklin 1799.
4. Wilson et al. 1993.
5. Simonson 1989.
6. Dijksterhuis 2004.
7. Hsee 1999.
8. Simonson 1989.
9. Hsee et al. 1999.
10. Thompson & Norton 2011.
11. Rozin, Millman & Nemeroff 1986.
12. 可参见：Thompson, Hamilton & Rust 2005。
13. Tversky & Shafir 1992, p.305.
14. 或至少参与者是这样认为的；可参见：Simonson 1990。
15. Thompson & Norton 2011.
16. Morabia 2006.
17. Simonson & Nye 1992, p.442。我们调整了美元数额，大体上把通货膨胀到出版的这段日期纳入了考虑范围。
18. Arkes & Ayton 1999.
19. McAfee, Mialon, & Mialon 2010观点类似。

第十五章
理性之光明面

在电影《十二怒汉》（*12 Angry Men*）的开头（先“剧透”一下），有个男孩受到指控，说他捅死了父亲，他的处境很危险：陪审团议事室里用于给他定罪的论据堆积成山。一个目击者声称见到这个男孩捅死了父亲；另一个目击者则称听到了打斗声并看到男孩逃出了公寓；这个男孩的不在场证明也说不通；他有动机，而且总是喜欢使用暴力，这些在档案上都有记录。群体极化现象正在酝酿中，蠢蠢欲动，试图说服陪审团将该男孩送到死刑电椅上去。但是，有一位陪审员与他人不同，他没有那么笃定。虽然该陪审员认为被告并非无辜，但也不能十分确定他有罪。他“只想探讨一下”。当被极力要求给出论据时，他用了一个没什么说服力的论据开头：证明男孩有罪的证据太合理了，这种合理性不得不让人产生怀疑。不出所料，这并没有动摇任何一个陪审员。不过，从那开始，这个陪审员渐入佳境，不断找到该公诉案中的漏洞。

他揭露了归罪论据的不一致性。一个目击者声称，在高铁列车驶过时，自己透过车窗从这条街的对面看到了这场谋杀。另一个目击者声称听到了男孩威胁父亲——“我要杀了你！”——然后几秒后就听到身体倒地的声音。但是列车驶过的声音震耳欲聋，第二个目击者是如何听见男孩和父亲的争吵声以及身体倒地声的呢？

其他陪审员的论据出现了更多的不一致。刀上找不到该男孩的指纹。但四号陪审员认为这不是问题：因为这个男孩是个冷血的谋杀犯，所以他还擦了刀，上面沾着其父亲的血。谋杀事件过去三个小时后，被告在回家的路上被警察抓住了。他为什么要回到犯罪现场呢？四号陪审

员答道："他杀了父亲后跑出来时非常恐慌，后来，当他最终冷静下来时，发现自己把刀落在那儿了。"但是，那个较多疑的陪审员接着又指出，之前的说法是"他很镇静地处理了这把刀，所以刀上没有指纹"，现在又是另一种说法，那么对这种不一致性做何解释。

解决这种不一致性可以采取双重标准。这个男孩来自贫民窟，十号陪审员清楚，那里的人是"天生的谎话精"，"他们说的任何一句话都不能相信"。然而，对于那个声称看到男孩犯罪的目击者的证词，他却坦然接受了，即便正如我们的十号陪审员指出的，"她也是贫民窟的一员"。

有时，这种不一致性一目了然，清晰可见，甚至都不用指出来。三号陪审员竭尽全力维护有罪判决，很大程度上是因为目击者声称听到打斗声并看到男孩随后立马离开了公寓，陪审员相信这个人的证词。然而，目击者声称自己在15秒之内起床并横穿了整栋建筑，想一想，对这个跛脚的老爷子来说，这可能吗？但是三号陪审员认为这不是问题：目击者肯定把时间估错了，毕竟，"他是个老爷子，经常稀里糊涂的，怎能事事记得准确呢？"

这些不一致性慢慢动摇了大多数陪审员，让他们产生了合理怀疑，但还不是所有人。十号陪审员对这些理性思考仍无动于衷。最终，其他陪审员一定是让他感到过意不去了，他只好变得宽厚些。尽管如此，论证还是发挥了最大作用。是论证让公诉案中的漏洞浮出水面，是论证突出显示了双重标准，是论证披露了这些不一致性，是论证发掘出陪审员心中的疑问。在《十二怒汉》这部电影中，论证救了一个男孩的性命。注

论证被低估了

在第十一章至第十四章中，我们已经强调了理性“阴暗”的方面，这些方面一开始就带来了一个谜团：理性是有偏见的；理性是懒惰的；理性让我们相信疯狂的想法并做出愚蠢的事情。如果把理性放到进化和互动视角中看，这些特点便有了意义：自我中心偏见在说服他人时很有用；懒惰在来回的互动中比较高效；若理性的使用范围超出合适的论证语境，可能会让人产生疯狂的想法。我们也已经反复强调过，所有的努力都是为了做到最好，也就是说，在合适的语境下，就应该把理性的这些特点变成分配认知劳动力的有效方式。

关键是，论证推理的这种辩护取决于人们评估他人理由的方式：人们必须认识到哲学家尤尔根·哈贝马斯（Jürgen Habermas）所谓“更好论证具有非强迫的力量”。[注]若接受足够充分的论据，就必须否决微弱的论据，即便那意味着完全改变看法。在第十二章中，我们摆出证据，证明人们擅长评估他人的论据，但是实验中的目标人群仍是单个推理者，这些推理者要评估的只是单一论据，与论证交流的互动有很大差别。

现在是时候兑现承诺了，那就是展示理性在合适的互动语境下的作用。理性让人改变彼此的想法，这样最终就能形成更好的观点、做出更好的决定。

20多年来，心理学家和教育研究员戴夫·莫什曼（Dave Moshman）一直让其学生处理沃森四卡选择任务（见图11.1），先是单独进行，然后以小组形式进行。虽然个人表现还是一如既往的差，猜对的人只有15%左右，但小组讨论的效果令人惊奇。半数以上的小组都回答正确。当莫什曼与莫里·盖尔（Molly Geil）搭档开展受控实验时，小组的正确率达到了80%。

有许多文章指出了沃森四卡选择任务中参与者令人遗憾的表现，如果不挨篇读那些文章的话，可能很难意识到实验结果的惊人之处。在该任务的标准模式中，没有任何参与者的正确率接近80%。就读于美国顶

尖大学的大学生若独自解决任务，则正确率甚至达不到20%或25%。[注]有参与者努力寻找正确回答，但最终还是彻底失败了。[注]

莫什曼和盖尔所做的，相当于让短跑选手汇聚一堂，在五秒内共同完成百米赛跑。

你可能会以为，这一惊人的实验结果会引起心理学家的注意。根本不是这样，该结果完全被忽视了。也许没有人真正知道该如何利用这个结果。唯一关注莫什曼和盖尔实验结果的研究者，还是一些自身理论被削弱的人。鉴于许多心理学实验都存在不能重复的可疑倾向[注]，他们要求用公正的态度重做这些实验。[注]小组讨论效果的改进虽然不能总像原实验那样激动人心，但被证明相当稳健。[注]运用到其他任务（如我们在第十二章介绍的保罗和琳达问题）时，同样效果良好。[注]你可以与朋友、同事或学生一起尝试一下该实验，效果屡试不爽。

有研究员表示怀疑，也认为论证与表现的改善没有多大关系。他们认为，小组成员并没有关注彼此论据的内容，而是依赖肤浅的归因来决定该采纳哪个回答。也许人们只是听取了最自信的小组成员的意见。[注]这一解释有点道理：不论是好是坏，信心会是促成小组讨论结果的决定性因素。[注]

然而，这一低水平的解释对小组讨论推理任务时的情况表述欠佳。看到讨论记录会发现，有些人观点占上风，显然并不只是因为他们用自信的口吻说了“我知道那是事实”，还因为他们给出了一个接一个的论据。[注]我们也清楚，若单个参与者的答案正确，即便整个小组成员的答案都是错误的，即便他起初没有那么自信，也能说服其他成员相信他的答案正确。[注]

严格意义上讲，若不可能证明给定答案正确、其他答案错误，要如何进行论据交流呢？若论证缺乏这种论证的力量，其他因素对说服他人

或评价他人观点则会起到更加重要的作用，如谁看起来更有能力、有多少人支持某给定观点。但是，即便对于没有明确解决方案的问题，小组整体表现通常还是要比小组普通成员好。在某些情况下，小组整体表现甚至比小组中最优秀个体的表现还好。㊟顺便提一下，即便小组成员共同的答案不如小组中最优秀个体的答案，用小组共同的答案还是更好，除非一开始就明确知道谁的答案最妙。

群智何时有用，何时失灵

怀疑群体效能不完全是错误的。若不运用论证，团队表现会令人失望。100年前，农艺学研究员马克斯米连·瑞吉尔曼（Maximilien Ringelmann）注意到一个奇怪的模式：若拖拉机、马和人类一起执行任务，似乎效率要低一些。㊟譬如，若一起执行任务，则人类推动马车时就不用使那么大力。

由于机器、动物和人类一起用力时整体的绩效下降了，瑞吉尔曼便把大部分责任归咎于协调问题，他认为，因为没有同时用力，所以在任何给定时刻施加的总力量就减少了。不过，通过观察囚犯推动磨粉机，他还注意到，对人类而言，激励会是个重要因素："成果平淡无奇，因为只过了一会儿，出于对旁边人的信任，每个人都满足于顺应曲柄的移动方向，甚至有时还被曲柄带着走。"㊟几十年之后，社会心理学家将证明，这类激励因素往往是团队表现不佳的罪魁祸首，并称这种现象为"社会惰化"。㊟

团队不仅在需要付出体力的问题上表现得令人失望，在各种认知问题上也是如此。头脑风暴就是个典型例子。总的来说，团队头脑风暴并没有什么作用。在典型的头脑风暴期间，不允许参与者表达对他人看法的苛责之词，这样大家便都能自由地发表看法，甚至可以是疯狂的想

法。其实这样做没什么用：与汇聚起来的个人独立看法相比，团队头脑风暴出来的想法通常更少、更糟糕。[注]相对而言，若告诉人们“大多数的研究都建议他们辩论甚至批评彼此的想法”，则会让他们想出更多的主意。[注]

许多领域团队的表现都让人失望，导致成功的论证更令人瞩目。若人们进行论证，即便有关看似无趣的数学或逻辑任务，也不会出现社会惰化或认知中断，而是增加其在对话语境下的积极性。他们回应彼此的论据，并在此基础上提出新的论据。许多伟大的思想家也已注意到激烈辩论对刺激其智力的重要性。下面是蒙田（Montaigne）的看法。

> 书籍研究日渐衰弱，变得无力，不再是热点，而对话却能当即教导人、锻炼人。如果与我对话的人意志坚定，堪称一名雄辩者，则会攻击我的两翼，将其逐个击破；他的想象力会激起我的想象力；忌妒、荣誉和竞争，激励我做出一些超出能力范围的事；另外，默许是对话中令人生厌的行为。[注]

论证让人对各色任务做出更好的回答。不过，到目前为止，审查结果都来源于严格控制的环境下所进行的实验，实验中的参与者之前都没有见过面，今后也不会见面。在真实的世界里，情况就不同了。问题会变得非常困难，而且可能没有彻底的、令人满意的解决方案。科学家找寻着主宰宇宙的法则。政客试图在存在严重分歧与敌对情绪的议会里通过法律。法官找寻途径对合法但相互冲突的利益给予应有的尊重。在这些情况下，个人偏好与亲和力会造成干扰，甚至获得了优先权。坚定的主张与价值观受到攻击，但得到坚定的捍卫。论证在应付更复杂的问题、战胜情感的执念时，仍会发挥积极的作用吗？

如何做出更好的预测

预测很难，尤其是关于未来的预测。讽刺的是，这一常见格言也说明了学习过往的难度，这是因为过往已经被归因了，以这种或那种形式，归因到每个人及其表亲，从孔子到尤吉·贝拉(Yogi Berra)。[注]从事专业政治判断的专家菲利普·泰特罗克（Phillip Tetlock）想要弄清楚做出合理的政治预言到底有多难。[注]于是，他于20世纪80年代末期招募了300个政治专家，其中许多专家都有哲学博士学位，还有多年的从政经验，他让这些专家各司其职：做出对政治事件的预言。15年匆匆而逝，现在可以将这些预言与实际结果相比较了。专家表现得如何呢？非常糟糕。他们比玩飞镖的黑猩猩（代表随机答案）好不到哪去，而且轻易就被基于现有数据的简单统计推断超越了。[注]预测很难。[注]

从某种程度上来看，专家面临的主要问题并不是他们的预测不够精准，而是他们尚未意识到自身的预测不够精准。世界很复杂，就连专家都遭遇到了严格的认知限制——就其所能获取并处理的信息数量而言。但是这并没有阻止他们做出极端的预测：专家经常说某给定事件几乎肯定会发生或不会发生。[注]专家在其津津乐道的理论的影响下，预言未来时过于自信了。

遵循第十四章中展现的极化和过分自信实验，泰特罗克发现，推理偏见造成了这些极端，通常还促成了错误的预言。他观察到，当专家做出预言时，“不太容易认真看待他人的观点”，[注]而且“其一边倒的辩护也在助长过分自信的气焰”。[注]

理性也让专家修改自身观点的方式发生了扭曲。若正如其偏爱的理论预期的那样，事件真的发生了，专家则会变得更有自信。但若并非如此，我们的政治专家就会变成“找借口专家”。这场战争蓄势待发。历史的小插曲让他们的预言没能成真。无论怎样，政治太复杂了。[注]专家寻找借口时已经十分娴熟了，因此即便事实证明他们错了，对其理论正确性的笃定还是只增不减。[注]若专家能够进行协商，情况能变得更好吗？

与“冷战”（Cold War）时期相比，若“恐怖均衡”几乎不能预防原子战争的全面爆发，对我们而言，预言则比以往任何时候都重要。美国空军作为其中一个参与者，期待能更好地预测核战争的影响。兰德公司（RAND Corporation）由此而来。要改善其预测，可以平衡几个专家的观点，这样既简单又有效。不过兰德公司的两个研究员，诺曼·达尔基（Norman Dalkey）和奥拉夫·赫尔默（Olaf Helmer）认为若让每个专家了解其他专家给出的回答（比如说通常性回答），他们可以做出更好的预测。专家基于这一信息做出新的预测；实验人员平均这些预测，然后再一次提供给所有的参与者，让参与者做出另一个预测；以此类推，进行几个轮回。

这一反复平均的技巧，被称为德尔菲法（Delphi），其初次使用是为了弄清楚俄罗斯人为减少美国工业基地四分之三的产量，会投放多少颗炸弹。[注]幸好，在这一具体的案例中，达尔基和赫尔默还未发现其方法是否做出了精准的预测，但是自20世纪50年代以来，许多研究已经表明，德尔菲法能改善各种预言，不论是捍卫立场，还是医学诊断。[注]

与面对面的讨论相比，德尔菲法有好几个优势。在面对面的讨论中，提供最好的预测可能不如取悦上司或达成共识那么重要。面对面的讨论也需要让一群忙碌的专家于同一时间待在同一个房间里，安排起来总是不那么容易。但是，德尔菲法的匿名问卷调查解决了所有的问题。

然而，如果论证理论正确的话，德尔菲法的初期形式就与改善预言的主要方式失之交臂：交流理由。若佐伊（Zoe）认为意大利有80%的概率赢得下一届的世界杯，迈克尔却告诉她只有20%的概率，那么佐伊应该怎么做呢？平均这两个观点，然后把概率调整到50∶50吗？通常，若人们处于这一情境，会选择相信其他观点。譬如，佐伊可以将获胜概率定在60%。毕竟，她清楚自己为何认为意大利获胜概率更大，却不知道迈克尔所持观点的理由。[注]若她知道迈克尔的理由，知道迈克尔为何认为意大利获胜概率低，可能更倾向于改变自身的看法。

在德尔菲法中，专家不只是接受其他预测的均值，还能知道支撑他人预测的理由。吉恩·罗维（Gene Rowe）和乔治·莱特（George Wright）探究了理由差异对预测的影响。[注]事实上，这些不同的理由并没有促使专家更频繁地改变自身看法。的确，若只看到了均值，他们更可能坚持自身起初的看法。但是推理并没有失败，它只是在表现其识别力。

并非所有理由都是合理的。若迈克尔告诉佐伊，他认为意大利只有20%的概率获胜，是基于一只章鱼的预测，[注]佐伊肯定会迫不及待地更新自己的预测。但若迈克尔告诉佐伊，他听到内部消息，说意大利队核心前卫的身体每况愈下，佐伊可能会不假思索地认同迈克尔的看法。在罗维和莱特的实验中，参与者正是如此做的。这些参与者并没有过多地改变自身看法，而是更加坚定自身立场。需要时，他们会朝着正确的方向改变自身看法，而且会做出更好的预言。[注]

提供一次性的理由汇总是向前的重要一步，却没能充分利用推理。当人们可以交流论据和反对论据时，推理便能在对话的往复中熠熠生辉。网上交流能让专家交流论据时保持一定距离，进而可以更好地预测。

菲利普·泰特罗克20年前研究了政治专家的判断，20年之后又与芭芭拉·米勒斯（Barbara Mellers）及其他同事一起，开展了一项更有野心的实验。[注]他们一共招募了1300多名参与者，让其对地缘政治进行预测。第一组参与者单独进行预测，不能获知其他参与者的预测，这样便可以保持其判断的独立性。第二组参与者也单独进行预测，但是如德尔菲法的早期版本一样，可以获知有关他人预测的统计学信息。

第三组参与者被分成几个小队，每队大约20人，让他们一起在网上讨论这些预测。与独立的预测者相比，他们几乎所有预测都更加准确，而且他们三分之二的预测甚至胜过第二组。很明显，论据的交流让他们

做出更好的预测。

为陪审团辩护

到目前为止，我们已看到，若允许参与者进行小组讨论，他们就能为各种任务提供更好的答案。我们还认识到，每组专家也都能利用论证来改善自身预测。不过，本章开始提到的陪审团议事室（虽然是虚构的）中的讨论，并不是一项严格控制的实验或专家间的网聊。陪审团议事室是典型的面对面环境：暴脾气可能会加剧、寻求共识非常重要、错误不会被消除。陪审员是典型的“非”专家人士，对法律只有一点贫乏的了解。在《十二怒汉》这部电影中，他们的讨论充斥着偏见，他们的决定也被许多其他偏见搅得混乱不堪。

陪审员常常会在听证后形成对判决的错误直觉。因为有自我中心偏见，单个推理不可能帮助纠正最初的直觉。审议可能变成群体极化，会放大而非纠正共同的偏见。鉴于这些局限性，《十二怒汉》的确看起来比较乐观，是对论证力量的启迪与美好颂歌。也许我们应该遵循卡斯·桑斯坦（Cass Sunstein）及其同事的建议：把陪审员的一些决策拿给“研究这一主题的专家”看。[注]

在给人定罪的陪审团前，我们应该考虑替代性选择：法官和专家。爱德华·柯克爵士（Sir Edward Coke）无疑就是这样一个专家。他是英国16世纪末到17世纪初的法学家，还“可能是史上最博学的普通法律师”。[注]他的观点可以“追溯到无数的中世纪文本中，这些文本大多是手稿卷轴，他孜孜不倦地精读了这些卷轴”。不过，他精读这些卷轴，并不是因为爱好古文物研究。柯克“显然希望找到符合其法律及政治上坚定主张的先例”，而且他“有时甚至曲解先例来支撑自己的判决”。威廉·布莱克斯通爵士（Sir William Blackstone）在其1766年发表的、有巨大影响力的《英国法律评注》（*Commentaries on the Laws of England*）中

警告道，法官的学识及智慧不能保证其观点就一定公正——他这时想到的可能就是柯克，他认为："要解决或协调事实性问题，若委托给任何一个法官，他们便有足够的空间表现偏袒和不公：要么妄下论断，却被事实证明并非如此，要么更巧妙地限制某些条件，延伸或扭曲另外一些条件，以与其他条件区分开来。"[注]

不论多么有能力的法官，都有自我中心偏见，他们会使用自身学识来捍卫已有观点，而非形成公正的判决，读者不应对此感到惊讶。很明显，陪审员也没能从自我中心偏见中豁免，但是，正如布莱克斯通所见，审议有可能会抵消每个陪审员的偏见。申斥了法官后，他继续说道："因此一群有能力、明智、正直的陪审员……会成为真理的最好调查员，社会正义最可靠的守护者。"200年后，美国最高法院的大法官哈里·布莱克蒙（Harry Blackmun）也用相似的措辞为陪审团辩护："要恰如其分地将群体常识应用到任何给定案例中，抵消各种偏见至关重要。"[注]

布莱克斯通对法官的悲观评价可能是正确的，[注]但是他认为陪审团审议能平衡陪审员偏见的看法也是正确的吗？这个问题不容易回答。我们无从了解陪审员交流的论据。我们最多从审议后的面访了解到，起初少数人的看法，有时能变成最终判决。至少可以判断，审议能改变陪审员的看法。[注]为了知道审议的更多影响，我们须依靠对模拟陪审团的研究。

20世纪80年代初期，里德·海斯蒂（Reid Hastie）、史蒂文·彭罗德（Steven Penrod）和南希·彭宁顿（Nancy Pennington）开展了一项对模拟陪审团十分重要的研究。[注]为了使其实验尽可能逼真，他们招募的都是那些做过陪审员的人。他们在把参与者送去审议前，给他们看三个小时真实审判的视频回放。在该审判中，被告被指控在互殴升级之后，把一个人捅死了。虽然这场谋杀已尘埃落定，但判决可以变动，若发现被告只是自卫杀人，则无罪，若发现被告早有预谋，则定为一级谋杀。虽

然究竟如何尚不可知，但许多法律专家认为，合适的判决是定为二级谋杀；真实的审判中也是这一判决，因此就有了以下的研究。

在该实验中，陪审员看完了三个小时的视频后，必须要说出赞成哪种判决。大多数人赞成过失杀人。只有四分之一的陪审员赞成二级谋杀。也就是说，大多数陪审员起初就站错了立场。不过，他们最终还是达成了最好的判决。即便在某些小组中，只有少数几个陪审员捍卫二级谋杀的判决，但往往就是这小部分人设法成功说服了整个陪审团。(注)审议让许多陪审团达成更好的判决。正如布莱克斯通预料的那样，审议让陪审员抵消各自的偏见。

的确，这恰恰是从事法学及心理学研究的学者菲比·埃尔斯沃斯（Phoebe Ellsworth）在重做海斯蒂及其同事的实验时观察到的。

> 个体陪审员往往会关注有利于其最初判决偏好的证词：关于先前两个男人对抗的证词，通常由支持谋杀判决的陪审员提出，而被害人在被害前立即用拳猛击被告的证词，通常由支持过失杀人或自卫的陪审员提出。这一倾向不是弱点，而是审议过程的益处——与起支撑作用的事实证据一起，提供了比较几种不同事件阐述的机会。(注)

《十二怒汉》这部电影也并不全是空想。就像在电影中一样，真实审判中个体陪审员也可能会有偏见，也会犯错，无疑也会维护靠不住的阐释。不过，审议能让陪审员更详尽、更客观地审查证据，弥补个人偏见并让陪审团达成更好的判决。埃尔斯沃斯用这些乐观的词来归纳其影评：“《十二怒汉》是一种理想，而且是可以实现的理想。”(注)

论证起作用

互动推理理论预言，人们应该擅长评估他人的理由，否决微弱无力的理由，并在理由足够充分时改变自身看法。尽管该预言看似微不足道，但就论证的力量而言，它与普遍的悲观情绪相悖。譬如，若要人们评估参与者独自或以小组形式处理逻辑任务的难易度，他们则会认为小组不会比个人做得更好。(注)即便是推理心理学家，这些本应更了解的人，也低估了小组的表现。

这里审查的结果证明这一悲观但普遍的论证观点是错误的。我们时常看到，当人们遇到合理的论据时，会改变自身看法。不论是处理逻辑任务，还是寻找开放式问题的新的解决方案，不论是专家还是外行，不论是思考地缘政治还是忖度该给出怎样的判决，在与同行辩论过该问题后，人们便能得出更好的结论。

在第十六章至第十八章中将看到更多的情境，看论证如何让合理的想法传播、如何让小组胜过个人，证明托马斯·杰斐逊（Thomas Jefferson）在写下面这段话时没有过分乐观。

> 真理很伟大，若任其发展，会战胜一切，它是错误旗鼓相当的对手，完全不畏冲突。除非人类的介入剥夺了真理的天然武器——自由论证与辩论，否则，若允许其自由地与错误博弈，后者将失去威慑力。(注)

---

1. 引自电影《十二怒汉》，该电影由西德尼·吕美特（Sidney Lumet）指导，原作者是雷金纳德·罗斯（Reginald Rose）(1957; Beverly Hills, CA: MGM, 2008), DVD。
2. Habermas 1975, p.108.
3. 在科斯米德斯的书中（Cosmides，1989），哈佛学生的正确率只有27%；在斯坦诺维奇和韦斯特的书中（Stanovich & West，2000），1/4的顶尖高校的学生正确率只有18%。
4. Johnson-Laird & Byrne 2002.
5. Open Science Collaboration 2015.
6. Oaksford & Chater 2003，p.305.

7. Mercier, Trouche, et al. 2015.

8. Trouche, Sander & Mercier 2014.

9. 可参考上述文献。

10. Koriat 2012; Levin & Druyan 1993.

11. Moshman & Geil 1998; Trognon 1993.

12. Trouche, Sander & Mercier 2014.

13. 可参见：Laughlin 2011。

14. 可参见：Kravitz & Martin 1986。

15. Ringelmann 1913, p.10，译者为克拉维茨与马丁（Kravitz & Martin，1986）。

16. Latané, Williams & Harkins 1979.

17. 例如，Mullen, Johnson & Salas 1991。

18. Nemeth et al. 2004。该文章表明，意见分歧只会增加形成观点的数量而非质量，但是有进一步的证据表明，形成观点的质量也会因为意见分歧而有所改善；可参见：Nemeth & Ormiston 2007; Smith 2008。欲获取关于意见分歧在小组讨论中起积极作用的更多证据，可参见：Schulz-Hardt et al. 2006。

19. Montaigne 1870, p.540.

20. Barry Popik,“Never Make Forecasts, Especially about the Future”,2010.10. 30, 网址：http://www.barrypopik.com/index.php/new_york_city/entry/never_make_forcasts_especially_abo

21. Tetlock 2005.

22. Tetlock 2005，p.51。

23. 可参见：Silver 2012。

24. Tetlock 2005, p.61.

25. Tetlock 2005, p.123。

26. Tetlock 2005, p.62。

27. Tetlock 2005, pp.121ff。

28. Tetlock 2005, p.128。

29. Dalkey & Helmer 1963, p.461.

30. 可参见：Linstone & Turoff 1976; Rowe & Wright 1999。

31. Yaniv & Kleinberger 2000.

32. Rowe & Wright 1996.

33. 真实的故事……“Paul the Octopus”，维基百科，网址：http://en.wikipedia.org/wiki/Paul_the_Octopus。

34. Liberman et al. 2012; Minson, Liberman & Ross 2011.

35. Mellers et al. 2014.

36. Sunstein, Kahneman & Schkade 1998, p.2129.

37. Caenegem 1987, p.14.

38. Blackstone 1979, p.380.

39. Ballew v. Georgia, 435 U. S. (1978), p.234, 埃尔斯沃斯（Ellsworth，1989）的书中引用过，全文请访问网址：http://www.law.cornell.edu/supct/html/historics/USSC_CR_0435_0223_ZO.html。

40. Braman 2009; Sunstein et al. 2007.

41. 可参见：Sandys & Dillehay 1995。

42. Hastie, Penrod & Pennington 1983.

43. Hastie, Penrod & Pennington 1983，p.60。

44. Ellsworth 1989, p.217.

45. Ellsworth 2003, p.1406.

46. Mercier, Trouche, et al. 2015.

47. Virginia Statute on Religious Freedom, 1786, Digital History，网址：http://www.digitalhistory.uh.edu/disp_textbook.cfm?smtID=3&psid=1357。

第五部分
狂野的理性

在第十一到十五章中，我们已经描绘了理性明明白白支持新颖的互动理论，而非标准唯理智理论的图景。但是，我们不是集中讨论过某些情况吗？在那些情况下，最可能观测到符合我们所偏爱理论的结果。我们运用的许多证据来源于实验，这些实验的研究对象是美国大学的心理学系学生。人类理性如何发挥作用，为何会从微小且可以说相当不具代表性的人类样本中进化而来，要得出关于这两个问题的结论难道没有一点冒险吗？在第十六章至第十八章中，我们扩大了调查的范围，并寻找证据证明可以在各种语境下发现理性的基本原理：如危地马拉偏远的玛雅群落、幼儿园操场、市民论坛、实验室会议等。

第十六章
人类理性通用吗

乌兹别克斯坦温差大，在-40~40℃波动，极少下雨，所以并不富饶。因为其气候非常不适宜居住且地理位置偏僻，所以封建制度得以一直维系，直到20世纪初期才被废除。[注]20世纪20年代，在乌兹别克斯坦成为苏联的一员后，莫斯科当局决定在全国各地开设几百所学校，那里的现代化进程就此开始。[注]实验心理学莫斯科研究所的研究员亚历山大·卢里亚（Alexander Luria）认为，这一变动提供了检验其想法的绝佳机会。卢里亚师从利维·维果斯基（Lev Vygotsky），认为人类是通过学习获得大部分认知技能的，其中就包括学校学习。卢里亚在乌兹别克斯坦招募了一些参与者，这些人刚开始上学，但在其他方面与附近村庄目不识丁的农民差不多。通过比较这两类人，卢里亚可以确定，即便刚开始上学，也会对认知产生影响。

卢里亚在1931年提出，“前往中亚的心理学远征”[注]的目标之一是研究逻辑推理。俄罗斯心理学家确信目不识丁的农民能用熟悉的材料进行推理。的确，若与外人辩论棉花生长情况，农民很可能会获胜。卢里亚找寻的有所不同，是能就某论据得出服务于自身目的的结论的能力，这与论据前提的真假无关，他认为未受教育的人是不具备这种本领的。卢里亚所用事例在逻辑上微不足道，参与者对其内容也不熟悉，所以才不得已评估论据本身的逻辑。

> 在极北区，白雪皑皑，所有的熊都是白色的。新地岛就在极北区。新地岛的熊是什么颜色的呢？[注]

受过一两年的正式教育后，乌兹别克人发现这个问题很简单。但是当问没有受过教育的农民时，他们绝大多数人似乎都很茫然，他们给出的答案是这样的，“熊还有不同的种类？”“我不知道，我见过黑熊，但我从没见过其他熊”。㊟

论证有多么不可思议

多样性给进化提供了蓬勃发展的空间。这只是因为在物种进化过程中，同一物种中每个个体的遗传特点有差异而已。然而，自然选择却吞噬了让其赖以生存的这种多样性。当某遗传特点允许其携带者同类繁殖时，该特点便在群体中传播，几代之后，每个个体都会携带该特点。但也有例外情况。某些特点是性别特异性，若在同一物种内没有普遍存在，其他特点就会转而变得更有利。但是，我们还没有弄清楚为何理性属于这种例外情况。若我们的观点正确，即理性是适应性结果，可以帮助解决协调问题、声誉管理问题及每个人都会遇到的沟通问题，那么它就应该是正常人类所共有的，而不只是为少数人所有，也不只是为男人或女人所有。

在每个个体中，理性若要正常发展，需要一些输入，对话、论据均可。不同文化和环境下输入可能会有所不同：从数量上说，某些文化强烈支持论证，而另一些文化中则略有压制；从质量上来说，不同的论证形式，可能获得赞同，也可能遭到反对。不过，无论是何种社会形态，都依赖多种多样的沟通，而其他物种则不会如此，而且这种依赖为理性的诞生造成了选择性压力。因此，若我们的分析方式正确，理性便不可能是制度的文化产物，诸如学校教育这样的制度才持续了几百年而已。

有人认为理性是有历史定位的文化发明，这一想法已成为社会科学领域的老生常谈。在卢里亚的探险之前，法国理论人类学家吕西安·利维-布鲁尔（Lucien　Lévy-Bruhl）描绘了一幅“原始思维”的图画，“即便

效仿没什么抽象概念的推理也是一种未开化状态”。[注]他和该领域的其他人都主张，生活在其他文化中的人可能也会推理，却基于完全不同的逻辑。有两种观点与我们捍卫的进化理论不相容：一种观点认为，理性是近期的历史发展元素；另一种观点认为，理性在不同文化间表现为完全不同的形式。

到目前为止，历史学、人类学及语言学等方面的证据都指向我们论据中一个有潜在“危险”的瑕疵：关注点都放在西方文化中的例子及实验上。正如一群研究不同文化的心理学家及人类学家最近所言，这些人是“不可思议”（WEIRD，由Western, Educated, Industrialized, Rich, Democratic这五个英文单词的首字母组成）群体，他们来自富强、民主、教育体系完善的西方工业化国家。这一首字母缩略词名副其实，因为该样本经常处于诸多可测人口统计变量的极端区域。譬如，到目前为止，心理学实验中最多的参与者是美国大学生，他们比没上大学的同龄人更加自我，[注]没上大学的同龄人则比上一代美国人更加自我，而上一代美国人则比同时期地球上任何人更加自我。[注]

“不可思议”群体都比较重视论证可能就是另一个怪异的特点，这是因为他们看到古希腊人在科学、政治及法律体系中都依赖论证。在大多数西方文化语境下，分歧都被视为人类互动的正常方面，应该使之系统化，而不是遭到压制，分歧也能产生积极影响。尤其是大学，应该鼓励辩论实践。

这样看来，难道将论证视为有益的不是文化偏见吗？难道理性——至少是我们所描述的具备辩护及论证功能的理性，不是一种文化习得性技能，而是进化的普遍特性吗？譬如说，其他文化中的人能成为出色的孤独推理者而非糟糕的辩者吗？或者说根本不能成为推理者吗？

怎样避免看起来像个傻瓜

得出“新地岛的熊都是白色的”这一结论似乎不可避免，所以只有轻信的人才会认为有人得不出这一结论。尽管如此，最近的研究证实了卢里亚的结果。他重做了实验，实验参与者是几个没受过教育的人，[注]而且其他实验也都提醒我们，即便是最理所当然的技巧，也需要通过教化获得。

设想，把三个硬币一个接一个地投进一个不透明容器中，然后让你找回这些硬币。难道会有人做不到吗？事实上，会有。只有学过数数的人才能找到。如果你出生并生活在毗拉哈（Pirahã，亚马孙河流域的一个小部落），要你找回这些硬币，你可能会不知所措，因为那里的人对数字一点概念都没有。彼得·戈登（Peter Gordon）所做实验的参与者就是毗拉哈人，实验很简单，却有三分之二的人不知从何下手，而且若增加硬币的数量，他们的表现还会迅速恶化。[注]毗拉哈人完全没有想到这个美国人会让他们玩这么古怪的游戏。他们非常擅长做其他事，但真的不会数数。

所以，有些人只是因为不熟悉前提，才不会制造并评估论据吗？不，事实证明这个问题只是推理中的一个驱动力，或者更具体地说，是社交礼仪。在参与测试的所有人中，某些人——大概是参与者的三分之一——很容易给出正确答案。他们并没有独自形成一种新的认知能力。这些人只是更喜欢玩实验者提供的这个游戏而已。

为什么大家都不愿意回答这么简单的问题，如“新地岛的熊是什么颜色的”呢？在小范围的人群中，人们对自己的断言非常小心谨慎，只有当理由足够充分时才会表明立场（在这一点上，和专家大不相同）。[注]显然，若陌生人讲的故事很古怪，说很远的地方生活着颜色很奇怪的熊，则并不符合这些条件。只有傻瓜才敢做出这样的声明，因为对此不能进行合理辩护。正如在第十四章中所见，人们试图避免做一些辩护不了的事情。这恰恰是在下面的交流中所发生的事情，在实验人员询问未受过正式教育的乌兹别克成人白熊问题时便可捕捉到，交流如下。如果

一个人60岁或80岁了，曾经见过白熊，还讲述过这事儿，则他可以被信任，但是我从未见过白熊，因此我说不准。我说完了。那些见过白熊的人可以判断，那些没见过的人什么也不能讲。

这时，年轻的乌兹别克人主动答道："从你们的话中可以推断出那里的熊都是白色的。"但是那位年长者总结道："公鸡知道如何打鸣，因而以打鸣为业；我的所知所言，也不过如此！"[注]他们都知道实验者的意图，但只有这个青年愿意在这个古怪的游戏中赌上自己的信誉。

为了让人们能更加坦然地面对不熟悉的前提，心理学家保罗·哈里斯（Paul Harris）与同事给未受过教育的参与者提供了一个更加丰富的语境。这回，问题的语境不在遥远但真实的地方，而在遥远的星球，除此之外，其他方面都与卢里亚设置的相似。在这种情况下，人们做出假设时就不会有太多顾虑，更容易做出逻辑性回答。[注]若我们在学校学过玩推理游戏，即从任意前提出发，最终得出古怪的结论，[注]则不需要接受正式教育便能掌握基本的推理技巧。人类只要大脑没有障碍，就能制造论据并评估论据。但是，他们辩论或遵守这些技巧是为了私用吗？

东亚的论证

2011年3月11日，强烈的地震肆虐了日本东北部地区。福岛第一核电站遭到严重破坏，这是自1986年苏联的切尔诺贝利核电站事故以来，世界上最糟糕的核事故。鉴于人们发现本应保障核电站安全的监管过程存在明显漏洞，日本政界难辞其咎，因而如坐针毡。

"核电村"（该昵称恰如其分）是被指控者之一。[注]在那里，私企和监管机构一团和气。这就是问题所在。日本学者铃木武（Takeshi Suzuki）认为，这个村庄就像被言灵（kotodama，对语言神秘力量的一种信仰）控制了。[注]言灵禁止发表能引发争论的言论，"充当阻碍论证

的社会一员”。更为普遍的是，因为面临维持社会和谐的强大压力，“日本人并没有接受训练进行辩论和推理”。日本文化不允许“核电村”村民讨论核能的危险性，这让他们构造了一个“核电站的安全神话”。[注]

引用卡尔·贝克尔（Carl Becker）一篇文章的标题，铃木武不是第一个谴责“在远东缺乏论证和辩论”的人。[注]贝克尔将这一缺漏的责任归咎于盛行于东方文化中且由来已久的文化准则，并概述为至理名言“雄辩……只不过是通向地狱的业力”（日本临济宗语录）。其他作者认为，在处理逻辑及论据时，东亚语言较为匮乏。通过大范围比较东方文化，中村元（Hajime　Nakamura）注意到，中文和日语本身就存在含糊性，这一含糊性为“逻辑概念表达”设置了障碍。[注]

东亚文化因谴责论证经常被单独挑出来作为好的案例来分析，用来支撑我们的断言，即论证的某些特点是广泛适用的。东亚文化里的人们会辩论并从中获利吗？鉴于其语言存在所谓的含糊性，东亚人可以选择合理辩论吗？答案非常肯定。东亚语言并没有阻碍逻辑与修辞的发展。公元前5世纪，中国的墨家学派创造了与西方的经院学派同样复杂、难懂的逻辑体系。200年后，伟大的韩非子创造了非常丰富的修辞学，罗马还没有修辞学者能与其“心理学分析的微妙”相媲美。[注]

那么关于所谓反对论证的文化禁忌呢，因为任性而违反了社会和谐要受到斥责吗？讽刺的是，对于各自反对雄辩和论证的宣言，东亚的知识分子从未停止对垒。孔子的《论语》中说道“君子欲讷于言而敏于行”。但这并没有阻碍儒家学者写出一部又一部的专著“回应对手的批驳”并“与竞争对手的学说进行哲学辩论”。[注]在道教的历史中同样到处都可以找到论据，尽管《道德经》断言“善者不辩，辩者不善”。[注]

申斥论证并没有阻止东亚的知识分子辩论。这没什么可惊讶的。辩论是知识分子用来谋生的方式。其他人就更倾向于尊重已确立的准则并受到智者的嘲笑吗？很难说。最近的历史研究披露，即便像日本这种社

会管控严格的国家都“不是贝克尔等人描绘的‘相对和平的’社会，而是一个在过去300年间以大量的意识形态冲突及肢体冲突（后者更频繁）为特征的国家，而且其争端经常以辩论的形式演示并记录下来”。[注]

即便日本早期惯例体现了对小组讨论好处的理解。7世纪初期，马宿王子（Prince Umayado）制定了一个新的体制。该体制的最后一个条目是这样的：

> 对重要事情的决定不应该由某个人单独做出，而应该是许多人在一起讨论。但小事就没多大关系，没有必要咨询许多人的意见。只有在讨论重大事情、担心会失败时，才应该与他人保持一致意见，以得出正确的结论。[注]

只有极少数有关东亚文化中论证的好处的实验证据，向西方文化中的实验结果靠拢。若要日本学生独自解决沃森四卡选择任务，与美国或欧洲同龄人相比，他们的表现也好不到哪里去。但当他们进行小组讨论时，思考着论证带来的好处，最终大多数小组都能得出正确的答案。[注]

日本文化可能轻蔑论证，但是这种轻蔑与福岛的事故没有关系。这只是监管失利的普通案例。监管机构因为收受了或多或少的贿赂，便服务于本应监管的私人利益。在日本，公务员若与企业合作，退休后便有望分到轻松又容易赚钱的工作，也就是官员空降。其他国家采取的方式或许有所不同，但也难逃厄运。我们不能因临济宗可悲且影响较广的政治惯例就归咎于它。

小型社会里的论证

几千年来，东西方文化都是依靠文字来制定复杂的修辞惯例、发展

以论证为中心的体系的。人类没有在这种错综复杂的文化语境下进化。我们祖先生活的环境与卢里亚实验中没有上过学的农民生活的环境更相近，甚至更接近于当代狩猎采集者生活的环境。即便我们认可每个人——受过正式教育或未受过正式教育的，都能进行推理，也不意味着每个人都会辩论。

“高贵的野蛮人”（noble savage）以及托马斯·霍布斯所说的“所有人对所有人的战争”（war of all against all），是对人类原始状态的两个经典描绘，两种情况都不适合论证。高贵的野蛮人不会有论证的驱动力。各自为政的人只会阻碍辩论的蓬勃发展。

这些观点与事实根本不沾边。我们的祖先既没有与他人和谐共处，也没有对他人发动无休止的战争。而且，论证在他们的社会生活中，可能至少发挥了与在我们的社会生活中同样重要的作用。在现代民主中，人们若要做出集体的决定，就会去投票站。我们的祖先则坐下来辩论——如果现今的小型社会能让我们了解过去的话，至少会如此。在全球大多数这种类型的社会中，若问题严重到威胁集体利益时，如生态危机、战争、共有资源保护，人们便会聚在一起辩论，商讨出大多数人都满意的解决方案。[注]

与典型的小型社会相比，现代社会即便最平等的社会形态都显得格外等级森严。狩猎采集社会里没有国王、没有将军、没有管理者来支使人。若某些成员影响力更大，不是因为他们有超自然或与生俱来的权威，而是因为他们为社会提供了服务，如狩猎、服役、讨论以达成集体明智的决定。当期待一系列严格控制时——例如在饱受战乱之苦的社会中，这一概括也站得住脚。

人类学家拿破仑·夏侬（Napoleon Chagnon）在亚马孙河流域的雅诺玛莫人（Yanomamö）部落里待了许多年，见证了许多冲突，所以把成年男性50%的死亡归因于暴力冲突。[注]不过，这种无休止的冲突状态下

并没有诞生森严的等级结构。若村长考巴瓦（Kobawä）想要大家都听他说，不能只是提高嗓音或拿着棍棒威胁村民。他必须依靠论证。正如夏侬指出的那样，“要是有人计划做一些可能很危险的事，（考巴瓦）只会指出危险”。不管他的立场如何，“他只是很婉转地施加影响，这样就不会冒犯他人”。注

在小型社会里，我们可以找到很多有关错综复杂的论证实践的例子。爱德温·哈钦斯（Edwin Hutchins）在特罗布里恩群岛（Trobriand Islands）发现了令人费解的法律论证及复杂的推理链。注麦克斯·葛拉克曼（Max Gluckman）在赞比亚的洛奇人（Lozi）中发现了一种专注于辩论的文化现象，其中，用以下这些丰富的词汇来描述某人论证的质量（或论证中质量的低下）：

kuyungula——没有直截了当地说问题

kunjongoloka——说话时偏离主题

kubulela siweko——没有理解就滔滔不绝地说

muyauluki——说话时没有触及问题重点的法官

siswasiwa——迷失在话语中的人

siyambutuki——随性的健谈者注

为证明论证在传统的小型社会中和在我们现代化的大型社会中同样有效，托马斯·卡斯特兰（Thomas Castelain）实地考察了危地马拉基切玛雅（K'iché Maya）的偏远群落，那里的人经营着自给农业，而且大部分时候，读写能力都不行，只会说本族语。注

在当地人的帮助下，卡斯特兰让基切的参与者独自处理所谓的守恒任务。解决这类任务需要参与者认识到物体在发生某些改变后给定属性不变。当前的情况是：实验人员给参与者展示两个可以盛同样容量水的

玻璃杯，然后问他们水在哪个玻璃杯中会升得更高。在第一个玻璃杯中，会丢入一团培乐多（Play-Doh）彩泥；在第二个玻璃杯中，同样大小的一团培乐多彩泥会被分成两半丢下去。

只有三分之一的参与者回答两个玻璃杯中的水会升得一样高。但是当小组讨论时，70%以上的人得到了正确答案。其实，只要一个参与者得出了正确答案，就几乎总能说服小组其他成员改变看法，这一现象恰恰能在“不可思议”群体文化中被反复观察到。

如何协调推理的进化、认知及人类学视角

有人认为，要想理解简单的理由，接受学校教育很有必要；抑或是认为在某些文化中，论证可能被压制了或是从未发展过。我们注意到了推理及论证广泛适用的特点，证明以上观念是错误的。推理及论证随处可见的同时，我们应该预测理性是否是一个进化了的模块，论据的产生及评估是否是理性的两个主要功能之一。

称理性为普遍性机制，并不是说它在任何地方都发挥作用，且方式恰好相同，也不是说特定社会只会以肤浅、表面的方式来影响理性。各种文化习俗和体系都用到了理性，但可能会阻碍理性发挥某些用处，增强其他用处，也会提供社会发展性认知工具，补足并拓展自然进化的能力。

譬如，在各种社会形态下，无论是私人争端，还是刑事案件，这些权利问题都有规则和体系来解决。不同文化间的人辩论这些问题的方式也大不相同。在许多小型社会里没有司法系统，当事人就算证明自己有理也没有律师的帮助，由长老、当地议会或政治领导人担任仲裁者或法官。尽管如此，体制化的认知警觉可能会在这些诉讼中发挥作用。一方面，可能会对沟通者的文化发展性有所警觉，如让他们发誓或接受严酷

的考验，人们认为这样做可以阻止沟通者撒谎。另一方面，对沟通者所沟通内容的警觉，通常只依靠常识来产生及评估论据。通常没有规则来规定可靠证据，也没有取证标准。相对而言，在有国家机构的较大型社会中，仲裁、诉讼和刑事司法都掌控在复杂的体制下，遵循一系列确切的规章条例；当然，也存在许多文化差异。

近代史上最知名的一场刑事审判的对象是美国足球明星O.J.辛普森（O. J. Simpson），该审判惊人地阐释了受到法律严格控制的论证与常识推理相违背的程度。辛普森被指控谋杀了前妻及前妻的一个朋友。由于辛普森的律师在为期11个月的审判中始终主张辛普森与谋杀没有半点关系，因此，1995年10月，他最终被无罪释放了。然而，两个受害者的家属却已对他提起“意外致死”民事诉讼。在这场民事审判中，辛普森败诉了，于1997年2月被判支付3 350万美元的赔偿金给两个死者，因为陪审团断定两起死亡是他造成的。

对不熟悉美国法律体系的人来说，这可能看起来明显没有条理：这起双重谋杀案已判辛普森无罪，可为何他还要支付赔偿金呢？主要还是因为刑事审判与民事审判的证据标准存在差异。刑事审判的证据标准是“排除合理怀疑”，辛普森的律师主张确有合理怀疑，然后便胜诉了。而民事审判的证据标准是“明确且有说服力”，民事当事人辩称有这种证据并完胜辛普森。即便如此，大多数人，即便是美国人，也认为凭直觉很难理解这一双重判决。但是，法律学者通过更抽象的推理，却能很容易找到理由支撑这两个不同判决：刑事审判（你可以判处某人监禁或死刑）和民事审判（危险的只是钱而已）都可以找到合理理由支撑不同的证据标准。在辛普森的案子中，法律论据很可能与常识推理冲突。尽管如此，如同所有论据，这些论据最终还是根植于对高阶理由的直觉。

通过论证可以弄明白事件及权利相冲突的解释，这十分符合互动推理理论及论证推理理论。并非所有文化发展性推理都会对这些理论进行明显阐释。在第九章中提到了拼图、谜语、悖论等智力游戏，在某些文

化中，这些用来给孤独的推理者带来乐趣。这要如何符合互动理论呢？某些论证性互动的例子也明显不符合该理论，如牛津大学辩论社（Oxford Union，建于1823年，目前是竞争性辩论的世界冠军）这样的辩论协会。在竞争性辩论中，两个及以上的小队试图以论据击败对方，为任意分配给他们的立场辩护，就受众没怎么留心的话题进行辩论，如“国家应该虚构历史来提升社会凝聚力吗？”“若有一剂药可以阻止你陷入爱河，作为一名18岁的少年，你应该服用吗？”[注]人们为何要费事地去辩护那些自身并不支持的、探讨无关之事的观点呢？为何要关注这种论据交流呢？

在第十章中，我们主张推理有一个双重论证功能：对于沟通者，可以用推理来产生论据以说服警觉的受众；对于受众，可以用推理来评估沟通者给出的论据，若论据合理就接受，若不合理就否决。孤独的推理者将推理拼图及谜题作为休闲活动，在这种情况下，不必拿出说服他人的论据，也不必评估他人给出的论据。在竞争性辩论中，拿出论据不是为了说服受众相信其结论的真实性，而是为了令陪审团信服其论证技巧。

尽管如此，上面这些有趣的案例也没有对论证推理理论形成更大的质疑，与消遣式性交、手淫及色情文学对“性交的主要功能是生殖”这一主张的质疑相比。一些进化考量应该帮助证明。

任何进化后的机制都能适应它所进化的环境状况，但若环境变了，则这些机制可能会失灵或产生非功能性影响。若吸入的二氧化碳量高于空气中的正常含量，则可能会导致窒息。若吸入过多的氧气，则会导致狂喜。一般观念认为哺乳动物的呼吸机制很好地适应了向身体输送氧气并排出二氧化碳的功能，虽然出现了这些反常的影响，但并没有挑战这种一般观念。

进化后的认知模块通常适用于处理给定范围的信息，并从中得出确

切的推论。我们把这样的范围称为认知模块的"适用范围"。[注]譬如，避蛇模块的适用范围是该环境中有蛇，读心模块的适用范围则是个体正在互动或可能会互动的那人的精神状态。触发这种认知模块的是其他模块提供的输入信息，如避蛇模块这种低阶知觉模块，又如为读心模块处理他人行为、言语及视觉表达等信息的模块。

触发模块运作的输入是不完美的。它们没能挑出手头所有属于模块适用范围的例子，这需要耗费时间和精力。因此，接受一定比例的失误不失为高效之举。实际上，触发大多数模块运作的是简单的诊断性标准，而非对各种因素的复杂考量（该考量可能需要其自身的认知机制，而且不管怎样都不会那么有效）。[注]触发器往往过分敏感，如触发忌妒或危险探测，但这种过分敏感性很可能是适应性的。

激活模块的输入信息的范围就是其实际范围。实际范围不可能恰好与合理域相一致。探测失误无可避免，会导致假阳性错误（错误警报）或假阴性错误。正如第十一章中所见，基于假阴性错误（如把蛇误认作木头）与假阳性错误（把木头误认作蛇）的相对成本，有效的模块可能会偏向犯成本较低的错误，而尽可能避免犯成本较高的错误。在这种情况下，模块合理域与实际域之间不一致其实是有好处的。

不同物种间的互动（如猎物与捕食者的互动，其中，一个物种的拟态伪装导致另一物种的假阳性错误，或一个物种的伪装导致另一物种的假阴性错误）可以利用认知模块合理域与实际域之间的不一致性。同一物种间的社会互动也能利用这种不一致性。文化的许多层面正是基于这种利用。譬如，相信超自然中介的存在，这一信念可能根植于人们过分探究中介性与意图性的倾向，以至把精神状态归因于太阳、山脉等非中介物，而且能在自然形态中看到虚构中介行为带来的影响。[注]文化观念、习俗和人工制品的发展多亏了模块实际范围与适用范围之间的不一致性。

推理的合理域是与他人之间的分歧，即思想观念的碰撞。推理旨在通过拿出令他人信服的论据或评估他人的论据来减少这种分歧，而且可能还要使自己信服。我们在第九章中已经论证过，推理的实际域，即触发其运作的输入类型，便是对思想观念碰撞的探测。在与他人讨论时，会出现很多思想观念的碰撞。这没什么可惊讶的；人与人之间的冲突会触发很多思想观念的碰撞。在这种情况下，推理的合理域与实际域便重叠了。

但是，推理的合理域与实际域并非完全重叠。典型的假阴性错误是：处于支配地位或隶属大多数群体的人可能很少注意到下属或少数人的观点，也未能发现分歧所在。当然也有假阳性错误，其表现为：我们读古代辩论术或观看电视辩论时，思想观念的碰撞要么发生在我们不能与之进行互动的第三方之间，要么发生在个体心智内部。自身思想观念的碰撞可能是内部形成的悖论。若两个想法或可能的决定都有直觉支撑，而且直觉强度不相上下，我们则不能下定决心。个体内部思想观念的碰撞可能也包括所采用的人工制品：用于娱乐的、基于文化产生的智力题，或用来训练哲思能力的悖论。

为什么对某些人来说，解决谜题、思考悖论或观看他人辩论竟然很享受呢？这里再次涉及了进化论上的考量。对于许多进化机制，若要实现其功能，需要投入精力、时间和努力，生物可能感知不到所带来的好处，至少不足以激励它们继续努力。从生物学上看，性交的主要好处便是促进生殖，但是动物（包括人类）在过去的大部分时间里并没有意识到性交能繁殖后代。强烈的性欲和性快感已经进化，可以促使动物交配。

一旦性交伴随着欢乐的回馈，要获得或增加这些回馈可能不仅需要进行基本交配，还需要各种形式的性辅助活动，但到底能否促进成功繁殖犹未可知。此外，性交给人类带来的回馈可能是经济和社会效益，而且可能在各种有创意的性实践中得到发展。

推理是一种高投资的认知活动，也会间接地有益于健康。我们认为，推理最终会有利于人的健康，突破信任匮乏下的沟通限制，而且这样可以更好地对他人施加影响，并主动接受他人良好的影响。不过，这种想法目前还不足以促使人们对推理投入太多的精力。所以，我们认为，推理提供了一些享乐主义的回报。同样，这并非对所有人都适用。尽管如此，那些因社会地位或个人性情不愿辩论的人，观看他人辩论也可能会很享受，而且还可以运用推理来评估。

譬如，竞争性辩论因为劫持了推理得以开展，它将推理应用到并非其进化目的所指的领域。受众倾听他人的论据不是为了形成更有力的观点，而只是享受观看竞赛、处理非常充分的论据所带来的这种纯粹的快乐。辩论者拿出论据不是为了说服受众，而是为了享受并展示其推理技巧。

古时的哲学争论、博主间就高深话题进行的论战及正式辩论不属于推理的合理域，推理还没有进化到可以处理这种刺激，但是它们却属于推理的实际域，该实际域便可以解释其文化成功。

对推理的解读——即作为进化了的理性模块的功能，不必与历史及人类学兴趣——对不同文化中理由及论据运用的显著不同所表现出来的相冲突。

我们认为，这样其实就对此做出了更好的解释，并形成了可以验证的假设，这些假设不仅关于推理的种类，还关于其可变性。

初期的推理者

在成人推理者身上试验发现，推理的某些特点是普遍性的。发展及教育心理学研究的惊人结果表明，这些特点很早就显现了。理性的习得不靠教学机构，也不靠以理性为导向的文化环境——儿童自发地运用理

性来维护自身想法及行动，运用理性评估他人给出的理由。

先来看一个虚拟人物——塞古尔女伯爵（Countess of Ségur）的《索菲的不幸》（*The Misfortunes of Sophie*）中的女主人公索菲。此书是19世纪法国儿童文学中的畅销书。索菲十分好辩的性格虽激怒了母亲，但读者看得很高兴。下面举个例子。

> 索菲：妈妈，你为什么不让我单独去看石匠，却要我和你一起去呢？而且去了之后，你为什么总让我和你待在一起？
>
> 妈妈：因为石匠扔的石头和砖块会砸中你，而且这里的沙和石灰会让你滑倒，甚至是受伤。
>
> 索菲：哦！妈妈，我会非常小心的，沙子和石灰伤不到我的。
>
> 妈妈：你还小，你不知道。我是大人，我知道石灰会把你烧伤。
>
> 索菲：但是妈妈……
>
> 妈妈打断索菲，继续说道：不要跟我推理这么多，安静点儿。我比你更清楚什么会伤害你、什么不会。我不允许你没有我的陪同就到院子里去。
>
> 索菲低下了头，什么也不说了。但是她看起来闷闷不乐，自言自语道：我就要去，这让我开心，我要去。[注]

尽管索菲只有4岁，却能很好地理解妈妈给出的论据并予以回复。这是作者极尽想象虚构的事吗？不，女伯爵观察到她8岁的孩子也是如此。此外，心理学研究也证实，儿童很小就可以拿出论据，有的甚至在18个月大时就可以。3岁的孩子经常交流论据，并将之用于三分之一的争论中，要知道，他们的争论不胜枚举。[注]

儿童推理与成人推理的两个基本特点是相同的，一是对自己的理由

有自我中心偏见，二是对自己的理由评估标准较低。朱迪·邓恩（Judy Dunn）和佩妮·穆恩（Penny Munn）注意到，3岁的孩子对自己的论据会更有感觉，而且同与其对话者相比，这种可能性高达15倍。[注]雨果3岁的儿子克里斯托弗（Christopher）想要靠自己爬上高而陡的楼梯，以便声称自己是个大男孩儿。但几分钟过后，他爬累了，想要大人把他抱上去，便指出他只是个小男孩儿。这种不一致性似乎并没有让他苦恼。他曾经还声称自己太累了，不可能立刻就上床睡觉。很明显，儿童使用论据的方式有改进的空间。平心而论，尽管他们的父母有时也是这样做的，但没有过多考虑论据的一致性。

听到孩子自发地说出许多论据，父母并不会感到震惊，其实，像索菲的妈妈一样，他们甚至可能会觉得这一技能还有点让人恼火。但是孩子不仅可以运用推理来解释为何不应该上床睡觉，也可以运用推理来解释为何有权拿走兄弟姐妹的玩具。此外，孩子还关注他人给出的论据，并予以评估。

心理学家对父母与孩子的互动方式进行了许多观察，努力记下教养风格上的差异。谈到告诉孩子应该做什么或不应该做什么时，一些父母表示自己主要依靠威严。其他父母的做法则稍微合理一些，会向小孩解释应该上床睡觉、洗澡、捡起玩具、不要打搅姐姐等事情的原因。心理学家比较了孩子的不同特点，看教养风格差异是否会影响其认知及行为。

这些研究展现了运用理由的明显优势：合理的做法能让孩子“抵制住诱惑、对反社会行为感到愧疚、犯错后予以弥补、不自私、进行高水平的道德推理”。[注]看来，孩子似乎是基于自己的理由而给出理由，至少在某种程度上是如此。

这些研究面临的共同问题是只展示了关联性。父母若讲理，孩子便能努力让自己做得更好。但这并不是说理性的教养风格就能让孩子做得

更好。同样，这或许既能解释为何一些父母倾向于对其要求给出理由，又解释了为何其生养的孩子更擅长抵抗诱惑或对反社会行为更容易感到愧疚。即便消除了这一担忧，情况仍然如此，影响孩子态度的是理由的给出，而与理由的质量完全没关系。

为了更合理地测试孩子对论据质量的敏感性，在史蒂芬·伯纳德（Stéphane Bernard）及法布里斯·克莱蒙（Fabrice Clément）的协助下，我们借助孩子开展了一项简单的实验。[注]实验人员用摩比玩具（Playmobil toys）给其中一个小女孩安娜（Anna）讲了一个关于狗的故事，恰好她也有一只狗。故事是这样的：一天，狗走失了，小孩子都被叫来帮安娜找狗。为了帮助孩子找到狗，引入了两个新人物。第一个人指向左边并说道：“狗往这边跑了，因为我看到它往这个方向跑的。”这是个非常有说服力的论据。第二个人指向右边并说道：“狗往这边跑了，因为它就是往这个方向跑的。”这是个循环论证。即便是3岁的小孩子，都对论据的质量很敏感：他们被更有力的论据说服的可能性更大。而且，两岁的小孩子也已经开始显露出同样的技能了。[注]

守恒任务是大约一百年前让·皮亚杰设计来考查儿童对初级物理及数学的理解程度的，我们在本章前面说到的任务与之相似，其中，我们是将培乐多彩泥球扔进水里。你可以在网上找到许多关于守恒任务的视频，了解小孩子执行这项任务的情况。在其中一个视频中，率直的小女孩要面对狡诈的实验者。实验者先放了两块饼干在自己旁边，又放了一块在小女孩的旁边，然后问道：“你认为我们分得公平吗？”“不公平！”小女孩着重答道。接着心理学家把小女孩的饼干分成两半，然后问道：“现在公平吗？”“公平。”“为什么？”“因为我们都有两块饼干！”[注]

我们可以用守恒任务来哄骗小孩子，让他们相信双方的饼干同样多，还可以用守恒任务来研究认知发展的各个方面，其中包括社会互动对促进理解的重要作用。20世纪70年代，瑞士[注]、美国和英国的心理学

家都受到了皮亚杰的影响，开始让六七岁的孩子两人一组来处理守恒任务。为了增加趣味性，心理学家创建了尽可能多的小组，其中一个小孩能独自处理守恒任务（守恒者，conservers），而另一个小孩则不能（非守恒者，nonconservers）。守恒者非常有说服力，更有可能去说服非守恒者，而非守恒者说服他们的可能性则要低得多，前者的可能性是后者的三倍。[注]

鉴于守恒者有合理的论据，非守恒者改变了自己的看法。他们不会一味听从更有自信或更机智的同伴的意见，而是当与守恒者讨论其他话题如“做得最好的电视节目是哪个”时，有相同的可能性赢得辩论。[注]

一般来说，若让两个非守恒者一起讨论守恒任务，他们什么也讨论不出来，因为没有辩论的基础。不过，若任务失败的方式不同，讨论便能出真知。下面来看守恒任务的一个变体：孩子们看到，等量的水分别被倒进细而长的玻璃杯和矮而宽的玻璃杯中。法蒂玛（Fatima）认为在细而长的玻璃杯中的水更多，而玛利亚姆（Mariam）认为在矮而宽的玻璃杯中的水更多。法蒂玛试图解释时就是在向玛利亚姆展示其思维方式的错误性，玛利亚姆也是如此。他们通过论证便能逐渐意识到自己的错误，并领会正确答案“两个玻璃杯中的水一样多”。[注]

从很早开始，儿童就受到了正当理由的影响，不仅有大人的理由，还有同龄人的理由。事实上，在某些情况下，与对大人理由的关注度相比，他们似乎更关注同龄人的理由。当老师清楚地解释某数学问题的答案时，学生不加深究地相信他，不必过多关注老师的解释，因为无论怎样学生都愿意接受这个结论。相对而言，若学生内部意见不统一，在遇到合理理由时多半会改变自己的看法。所以，当学生真的改变了看法时，更可能是因为有理由的支撑。学生可以是彼此最好的老师。

教育工作者挖掘出论证在教学法方面的潜力。合作性学习研究于20世纪初期起步，20世纪70年代汇聚力量，目前已成为“教育研究史上最

成功的范例之一”。[注]成百上千的研究表明，当学生以小组形式讨论某任务时，通常能对材料形成更深层次的认识。[注]在教室中开展合作性学习不总是那么容易，一定会出现分歧，但是还没有达到冲突的程度。要让学生充分讨论某些事是需要时间的。尽管存在这些实际的阻碍，但是截至20世纪90年代，美国三分之二以上的中小学老师都依靠合作性学习来教学，总的来说效果较好。[注]

学会更好地辩论

声称理性的基本特点广泛适用并不是要否决文化差异。同样，关注论证早期发展性特点也不是要否决3岁孩子与成人之间论证的差异。若不需要学习产生及评估理由的基本技能，人们无疑会更擅长论证。

要更擅长论证，最基本的方式便是掌握说服听众的局部知识。譬如，随着对朋友电影品位的逐渐了解，你能更有效地说服朋友去看你觉得他们会喜欢的电影。人们也能认识到对给定受众适合用哪种论据。譬如，科学家了解权威给出的论据在学术性文章中的分量不大。把握辩论的时机也是至关重要的。与掌握辩论的方式至少同等重要的是，要弄清楚适合进行论证的语境。但是与获取新推理技能及更擅长推理相比，这些实际上都不能等同。

在第十二章中我们看到，在讨论开始时不费力去寻找最有力的论据通常也在情理之中。找到有力的论据比较费力，而且并不总是必要的；微弱无力的论据有时也可以说服听众。此外，对话方给出的反对论据能帮我们更好地理解对方立场，进而找到更合适的论据。随着对话的开展，人们给出论据时能更好地考虑听众的立场。

孩子让这种互动在表现形式上明显而夸张。当他们撒谎时，不擅长考虑对话方的立场，譬如无视妈妈根据面包屑的痕迹可以一探究竟这个

事实。同样，孩子的理由明显太以自我为中心了。当妈妈问及为何要拿走兄弟姐妹的玩具时，孩子惯常回答“因为我想要”。这样回答便是无视兄弟姐妹的立场，因为他们也想要这个玩具；同样无视妈妈的立场，因为妈妈已经弄清楚是谁偷吃了面包。孩子的论据很容易反驳：“他也想要”“他先拿到的”等。

当这个小孩较好地适应妈妈的反对论据，认同或用新论据反驳时，这一互动便标志着推理机制运转良好。然而，某些反对论据很容易就能想到，以至人们在不得不使用时会很愤怒。譬如，若孩子与母亲之间再次互动，母亲很可能就会恼怒：“我告诉过你，那个理由不合理！”母亲想要孩子明白，他给出的论据没有说服力，想要孩子要么立即承认错误，要么拿出一个更有说服力的论据。

在上述例子中，孩子要认识到两件事。第一件事是，对于其论据“因为我想要”，有反对论据让它没有说服力。第二件事是，若没有预想到这些反对论据，会产生不利的社会性后果，即母亲会恼怒。这可能需要一点迭代，但孩子或许就此学会更好地论证。

实验已经演示了人们是如何在更复杂的话题上学会预测反对论据的。在第十二章中，我们讲述了由迪安娜·库恩及其同事开展的一项研究，[注]其中，同龄人之间的讨论让青少年和成人都能就死刑拿出更有说服力的论据，而且辩证地看待论据的能力有所提升。这并不是说人们必须要改变自身看法，保持中立，而是说人们要能预测出对方的某些论据，并予以反驳。其中一个参与者，明显反对死刑，引用了一个论据来支撑立场：“我知道杀掉惯犯可以阻止谋杀链。”不过，他提及这个论据只是为了在下一句中予以反驳，指出终身监禁也能实现同样的目标。虽然该实验反映论据质量真的有所改善，但涉及的范围有限。辩论死刑只能让人们更擅长辩论死刑。为使人们的辩论能力有普遍性的提升，研究员和教育工作者经常借助其他方法，如教授批判性思维。这通常涉及许多（假定的）论证性谬论（如人身攻击谬论、滑坡谬论等）和认知偏见

（如自我中心偏见）带来的经验教训。总的来说，这类方案的影响较小。注若人们非常擅长发现他人持有的谬论和偏见，会发现以同样批判性的眼光审视自己要难得多。注

在很大程度上，如果学习推理就是学习预测反对论据，那么最好的解决方案可能就是让人们接触到更多的反对论据，进行更多的辩论。若某人的第一个论据很容易就被击垮，若遭到对方合理论据的强烈反驳，他可能不只是了解了这些特定论据的内容，更为普遍的是还认识到了反对论据带来的挑战。此外，可能还认识到做出某一预测是件好事，既可以避免提出站不住脚的观点，也可以避免提出很容易就会被击垮、会损害自身可信度的论据。

迪安娜·库恩和阿曼达·克罗威尔（Amanda Crowell）让学生进行小组辩论，看是否可以让他们推理得更好，即便他们可能从未碰到过讨论的话题。注研究者比较了两种干预手段的影响，二者都被严格执行：三个学年里每周开两次会，一次持续50分钟。其中一种干预手段是开一系列涵盖社会问题的哲思课程。该课程需要写许多论文，有时还要花整堂课来讨论。另一种干预手段是让学生必须就某些话题交流论据，如家庭教育或中国的计划生育政策。这种干预是让学生在与同龄人进行好几个论证环节的辩论之前，花时间打磨自己的论据。

每年年末，学生在写作考试中都会遇到新话题，必须就此选择自身立场并予以捍卫。有的学生上了更标准的哲思课程，在论文写作中积累了更丰富的经验，然而，拿出的论据却更简单。接受过论证训练的学生拿出了更复杂的论据，通常涵盖问题的正反两面。

论证似乎能提高人各个方面的推理能力。在某特定话题上遇到了反对论据，人们就学会了在其他语境下对其进行预测。人们在一定范围内可以通过普适性类推来预期各种反对论据，虽然论证推理理论不能准确预测这一范围，但比起唯理智理论，该理论更容易对这些发现做出解

释。在库恩及其同事的实验中，上标准哲思课程的学生尽管在个人推理方面接受了集中训练，能力却没怎么提升。相对而言，被要求相互辩论的人深谙论证的这种互动，反而能写出更好的论文。通过学习辩论，他们学会了独自进行更好的推理。

---

1. “Geography of Uzbekistan”，维基百科，网址：http://en.wikipedia.org/wiki/Geography_of_Uzbekistan; “Uzbekistan”，维基百科，网址：http://en.wikipedia.org/wiki/Uzbekistan。
2. 本章的第一部分在Mercier 2011a中可找到依据，第二部分在Mercier 2011b中可找到依据。
3. Luria 1934.
4. Luria 1976, p.107.
5. Luria 1976，pp.108-109。
6. Lévy-Bruhl 1910, p.22.
7. 从相对不怎么讲究的意义来看，术语“个人主义的”可以用于跨文化心理学，特别是用于东西文化的比较中。
8. Henrich, Heine & Norenzayan 2010.
9. Cole 1971; Scribner 1977.
10. Gordon 2004.
11. Blurton Jones & Konner 1976.
12. Luria 1976, p.109.
13. Dias, Roazzi & Harris 2005.
14. Dutilh Novaes 2015.
15. “Japan's Well-Placed Nuclear Power Advocates Swat Away Opponents”, NBC News, 2014.3.12, 网址：http://www.nbcnews. com/storyline/fukushima-anniversity/japans-well-placed-nuclear-power-advocates-swat-away-opponents-n50396。
16. Suzuki 2012.
17. 这三个引用都引自上述文章第179页；也可参见：铃木2008。
18. Becker 1986。下述引用也来自这篇文章。
19. Nakamura 1964, p.534.

20. Lloyd 2007, p.10.

21. Liu 1996.

22. Combs 2004.

23. Branham 1994, p.131.

24. “Prince Shotoku's Seventeen-Article Constitution”（Jushichijo Kenpo）, SaruDama, 网址：http://www.sarudama.com/japanese_history/jushichijokenpo.shtml。

25. Mercier, Deguchi, et al. 2016.

26. Boehm et al. 1996; Ostrom 1991.

27. Chagnon 1992。郑重声明，The Yanomamö不是狩猎采集者，而是园艺师。

28. Chagnon 1992。郑重声明，The Yanomamö不是狩猎采集者，而是园艺师，p.134。

29. Hutchins 1980.

30. Gluckman 1967, p.277.

31. Castelain et al. 2016.

32. 这些话题其实在杜伦大学2014年的公开辩论赛上就已经辩论过，网址：http://www.debate-motions.info/other-tournament-motions/442-durham-open-2014。

33. 欲了解合理域、实际域的概念及其对文化研究的关联性，可参见：Sperber 1994; Sperber & Hirschfeld 2004。

34. Gigerenzer 2007.

35. Barrett 2000; Boyer 2001.

36. Comtesse de Ségur, Les Malheurs de Sophie, 1858, 网址：http://fr. wikisource. org/wiki/Les_Malheurs_de_Sophie/3；可以看到我们的翻译。

37. Dunnn & Munn 1987.

38. Dunnn & Munn 1987，p.795。

39. Grusec & Goodnow 1994, p.5.

40. Mercier, Bernard & Clément 2014; 也可参见：Corriveau & Kurkul 2014; Koenig 2012。

41. Castelain, Bernard & Mercier这样主张。

42. 守恒任务，YouTube，2007.2.10，网址：http://www.youtube.com/watch?v=YtLEWVu8150。

43. Doise, Mugny & Perret-Clermont 1975; Doise & Mugny 1984; Perret-Clermont 1980.

44. Silverman & Geiringer 1973.

45. Miller & Brownell 1975.
46. Ames & Murray 1982.
47. Slavin 1996, p.43.
48. Slavin 1995.
49. Slavin 1995.
50. Kuhn, Shaw & Felton 1997.
51. Mercier et al. 即将出版；Willingham 2008。
52. Pronin, Gilovich & Ross 2004.
53. Kuhn & Crowell 2011.

第十七章
有关道德与政治话题的推理

假设你在参加一个关于道德判断的心理学实验。实验者让你坐在一个小房间里，然后填一张包含一些小故事的调查问卷。其中一个故事讲述了一个纪录片，某些片段比较可疑：影片中一些人声称并不知道接受采访时有人摄影。当你被问及是否赞同电影制片厂的决定，即无论如何都要上映这部电影时，你表达了强烈的反对意见。为什么你的反应如此激烈呢？或许是因为你听到了墨西哥移民（该群体风评不好，无须更多负面言辞来描述）对美国的抱怨声，或许是因为你对最近隐私频频受到侵犯表示担忧，又或许是因为房间里有十分难闻的气味。㊟心理学家十分擅长暗中操纵人的行为。为研究令人恶心的食物对道德评价有何影响，他们采用催眠术，让人脑海中出现恶心的厕所、溢出旧比萨盒和脏纸巾的垃圾桶。其实是使用了臭气喷剂，这之后，一些参与者在恶臭弥漫的环境下填写了问卷调查。与吸入纯净空气的人相比，闻到了令人不快的恶臭的人在表现道德判断时更加强烈。其道德判断不是受到理性的驱使，而是臭气喷剂的驱使。

其他让人讨厌的影响甚至会更加乱人心绪，正如以色列囚犯读了科学文献后感受到的那样。2011年，三名研究员指出了以色列法官在做出假释决定时的奇怪模式。㊟新的一天伊始，法官一般相对仁慈，会同意约三分之二的假释请求。到上午十点，这一比例会跌至零。到上午十点半的时候，会出现强烈的反弹，假释率会重回65%。几个小时后，假释率又会跌到零。到下午两点时，假释率会重新回升，60%以上的假释请求会得到同意。快下班时，假释率很快又会降到非常低的比例。

这一模式不能用理性来解释。到底为什么？因为法官休息了一下。上午十点左右，法官会吃些点心，下午一点则去吃午饭。中场休息让其情绪和精力如一天伊始时那样好。但这种刺激很快就消退了，因为同意假释请求要写更多报告，囚犯出狱的希望也会迅速降低。不知道囚犯团体是否会为此给审判室买个零食售卖机。我们只知道，法官从来不会把自己累了作为拒绝假释的理由。

目前，不论是逻辑任务、预测还是判决等，主要探讨的都是或多或少可以有正确回应的问题。然而，在某些情况下进行推理时，不确定什么是正确做法，甚至不确定是否有正确做法，如推理美学及道德问题时。

道德理性通常与其他理性大不相干。尽管如此，我们仍能从唯理智理论的道德领域里找到对拿物。唯理智理论的道德理性观念过分简单化了康德的立场，认为理性可以成为而且应该成为处理道德问题的最终仲裁者。人们应该运用理性形成合理的道德准则，并在采取行动或做出判断时与其保持一致。20世纪的大部分时间，让·皮亚杰和劳伦斯·科尔伯格（Lawrence Kohlberg）等道德心理学家都在采纳某种唯理智观念，假定更好地运用理性可以让人更遵守道德。

然而，通过对道德观的反思，保罗、克尔凯郭尔等思想家认为其也受到了情感及直觉的合理支配。他们也在心理学家中找到了盟友，如先前段落中开展独创性研究的实验者。

虽然在第十一章至第十五章中引用的案例都坚定地反对个人主义方式，道德领域却对此提出了新的挑战。或许在道德领域里，孤独的理性其实能战胜直觉，给孤独的推理者提供指导，得以做出更明智的决定。或者情况恰好相反，或许理性在道德领域十分无力，即便有合理的论据，也不能让人改变看法。

推理如何让我们表现不道德

乔纳森·海德（Jonathan Haidt）于2001年发表了一篇名为“情绪化的狗与其理性的尾巴”（The Emotional Dog and Its Rational Tail）的开拓性文章。㊟海德认为，推理在此只是“摇动狗尾巴”，在事后给出正当理由，掩盖暗中操控一切的直觉及情绪的痕迹。先前在这章提及的研究符合海德的理论，因为这些研究显示，道德判断不是受到理性的驱使，而是受到不相关因素的驱使，如恶臭及疲倦引发的坏情绪。但是海德进行了更深入的研究，认为理性没有让我们去做正确的事，而是为我们没做正确的事提供借口。

20世纪70年代，梅尔文·斯奈德（Melvin Snyder）及其同事开展了一项巧妙的实验，实验结果表明，学生为了不坐在残疾人旁边，会不惜使用最浮夸的借口。㊟实验人员告诉参与者要对古代喜剧进行评价。房间由隔板隔开，两边各有一块电视显示屏，会同时放映两部电影。每个电视显示屏前都有两把椅子，一把空着，另一把则被同伴占着，其实就是假装成另一个参与者的实验人员。虽然其中一个同伴没有可辨识的标记，但另一个同伴有重金属支架，表明他有运动性残障。

参与者得知，每个电视显示屏上会放映不同类型的电影。一个是打闹喜剧，另一个是悲伤的小丑喜剧。参与者更喜欢哪部电影呢？结果表明，不论是哪部电影，只要不坐在残疾人旁边就行。为了不坐在残疾人旁边，他们匆匆编造对古代喜剧的偏爱之词。

自此便出现类似的论证。譬如，男性参与者为拿到有泳装照的体育杂志不惜调整自身偏好：若这本杂志报道更多与体育相关的内容，那么体育报道便是决定性因素；若有更多专题文章，那么专题文章便是决定性因素。正如老掉牙的借口所言，“我是为了专题文章才去读《花花公子》（*Playboy*）杂志的”。㊟

哲学家艾瑞克·斯琪威茨吉贝尔（Eric Schwitzgebel）将此逻辑发展到极端，并观察专业道德推理者的行为，这些人的工作就是阅读、思考并谈论道德理性，也就是俗称的伦理学教授。结果表明，伦理学家经过道德反思，与其他哲学教授相比，去投票、付会议注册费、回复学生邮件或做出不粗鲁行为的可能性更低。(注)

这些例子为海德的模式提供了支撑，还展示了道德理性的微不足道，不论是让大学本科生躲避令其不适的人、让人看衣着暴露的模特，还是让伦理学教授不去投票。(注)在这些案例中，理性化的做法都不会造成进一步伤害。大学本科生对打闹喜剧新的热情既不会伤害自己，也不会伤害残疾人。但是对另一程度上违反道德的行为，可能需要用到更强有力的合理化，而且那些合理化本身便有其可怕之处。

伟大的推理者，糟糕的合理化

几年前，本书作者之一雨果受海德的邀请去弗吉尼亚大学分享我们的理念。去夏洛茨维尔（Charlottesville）旅行，若不去托马斯·杰斐逊的故居蒙蒂塞洛（Monticello），这段旅程便是不完整的。对这位出生于“小山沟”的开国元勋，我们还有待了解。他酷爱阅读，所用书籍曾堆满两间大屋子；他对法国启蒙思想家的仰慕深嵌在那些大理石半身像里，永不褪色；他的独创巧思呈现在他设计的巨型方尖座钟里；他的建筑智慧则孕育出这一新古典主义建筑杰作。

然而这座房子没有一个奇观能离开其建造者、在这5 000英亩（约30 351亩）的种植园里劳作的人——奴隶。种植园里有约200个奴隶。(注)若杰斐逊手头比较紧，又要买花哨的半身像，又要有其他琐碎的支出，就会把奴隶像私人财产一样卖掉。(注)奴隶若不听话，就会遭到鞭笞。(注)还会被卖到遥远的营房“以儆效尤”。(注)

杰斐逊的许多传记作者都指出，㊟在革命年代，杰斐逊作为弗吉尼亚的种植园主，这样的举动也无可厚非。不过，杰斐逊绝不是革命年代弗吉尼亚典型的种植园主。他支持全民普及教育，创办了一所重点大学，反对酷刑，构思了弗吉尼亚的宗教宽容政策，下面这些话便是他写的："我们认为这些真理不言而喻，所有人生来平等，造物主赋予了他们某些不可剥夺的权利，其中就包括生命权、自由权和追求幸福的权利。"下面就来谈谈认知的不一致性。

杰斐逊十分擅长推理，他成了海德模式最具戏剧性的例子。杰斐逊在思考奴隶制度该何去何从时，轻易就找到了理由来反对奴隶解放。

在《弗吉尼亚记事》（*Notes on the State of Virginia*）㊟一书中，杰斐逊表达了自己的担忧，他表示，解放奴隶只会导致"一个种族或另一个种族的灭绝"。他本可以点到为止，但他真的很想让自己的观点站稳脚跟，所以，"在这些政治性反对意见中"，他"又添加了物质性及道德性反对意见"。黑人与白人不能和谐共处，因为黑人的体格与灵魂都存在许多缺陷。"红与白的精致混合物——每一种或浓或淡的色彩都是一种热情的表达，不是比永恒单调的黑色——像不可揭去的黑色面纱一样完全遮盖了另一个民族的所有情感——更好吗？"黑人可能"更爱冒险"，但只是"想要未雨绸缪罢了"。他们的爱只是"一种急切的欲望"。"他们的悲伤是转瞬即逝的"。"他们的存在似乎只是为了分享更多感觉，而不是更多反思"。总的来说，"这种不幸的肤色差异，或者还有能力差异，是其获得解放的强有力阻碍"。

这是理性最糟糕的表现。这是公然的偏见，把"更巧妙的形式对称"这种最主观的评价，变成了"天性造就的真正差异"这种客观评价。科学家认为，只有傻瓜才会相信红毛猩猩更喜欢"黑人女性，而不喜欢自己的同类"。才思敏捷者不禁认为，在谈到黑人为了"那一丁点儿娱乐""一直坐到半夜"这个问题时，他们"似乎不怎么需要睡觉"，但是，鉴于"动物在身体处于休息状态、不思考时，一定更愿意睡觉"这一点，

黑人应该“会想要睡觉”。实际上，鉴于这种情况，这位修辞学大师就主张，现在做什么都无济于事，黑人只能最卑微地顺从，谁让他们没有飘逸的长发。[注]

事情究竟如何？是因为杰斐逊害怕种族战争，还是因为他持有种族主义观念，才让他反对立即解放奴隶的？[注]相对而言，杰斐逊更赞成教育黑人青年，让他们离开父母，返回非洲学习，这一长远计划有点不切实际。[注]但是，这只是因为他觉得黑人比白人卑劣，所以才特别担心不同种族间的平等互动［只有与情妇莎莉·海明斯（Sally Hemings）睡在一起时，他才不担心她的奴隶身份］。为什么奴隶获得解放后要被送往非洲那么远的地方呢？因为“解放黑奴后，为避免他们与白人通婚，只能出此下策”。[注]毕竟“与另一低劣种族的混合只会造成堕落，让人不爱国、不追求人性的优越，还天真地满足于此”。[注]

要弄清楚人们在遭到询问时为何会持有如此信念已经很难，重构死人的思维过程就更投机取巧了。然而，我们都知道，杰斐逊成为奴隶主并不是因为其种族主义信念，而是因为在继承种植园时就一起继承了种植园的奴隶，想必也继承了那时的普通奴隶主该有的态度。后来，杰斐逊还是接受了同辈知识分子的启蒙理念。只有才思敏捷的推理者才能协调这个巨大的矛盾，杰斐逊用一种可悲的方式接受这一挑战，这也是他留给后人的遗产。

显然，理性并没有表现得像我们期待的那样，助人传递更开明的观念、做出更公正的决定。杰斐逊智力超群，且熟知其时代的现状，有最崇高的理想，他本应通过推理，坚守正确的信条，维护公正的行为。然而，事实并非如此，理性提供了便捷的合理化，让他名正言顺地保留奴隶及财富。可悲的是，这些合理化并非毫无影响，而是让他成为“人类学美国学派种族主义伪科学的精神教父”。[注]

这类案例可能会促使我们将道德理性安全地锁住并扔掉钥匙。不

过，我们也应该想到，若理性的合理化力量是无穷无尽的，那么理性便不是无限制的。有时人们找不到托词，所以就没有任何正当理由不遵守道德，或老老实实地遵守道德。通过实验，我们洞悉了学生的心理，为达到不与残疾人坐在一起的目的，学生竟硬生生地编造出对某部电影的喜爱。在同一项研究中，另一组参与者则没有得到那个机会，因为两个电视显示屏放映的是同一部电影。要为不与残疾人坐在一起的想法辩护，这组参与者甚至没有机会表达虚假偏好。因此，他们为自己成功辩护的可能性较低。同样，与杰斐逊同时代的某些人也不能为自己保留奴隶的行为辩护。乔治·华盛顿（George Washington）解放了自己的奴隶，还在遗嘱中分给他们财产。本杰明·富兰克林在世的时候也解放了自己的奴隶。这样看来，人们实际上似乎不能“为自己想做的任何事情找到理由或编造理由”，[注]除非他像托马斯·杰斐逊一样机智。[注]

推理能改变人们的道德观念吗

我们目前描绘的道德推理，经过海德的重点强调，符合用互动理论进行无瑕疵推理的某一方面。理性不会认真评估某观念或决定的道德价值，而只是寻找作为托词的正当理由，为人们一直想做的事情辩护，不论其是否符合道德，这便是自我中心偏见在起作用。满足于浅薄的理由及不周密的合理化便是孤独的理性带来的另一隐患，会让人缺少对自身正当理由及论据的批判性检验。

不过，海德的理论还有另一分支，即“摇动另一只狗尾错觉”。正如论证推理理论可能预料的那样，“关于道德观念，我们希望成功反驳对手的论据可以使之改变看法”。海德认为，“这种信念就像以为用手拨动狗尾巴让其摇摆就能让狗很开心一样”。[注]换句话说，不论你的论据看起来多么有说服力，都不会改变他人对道德问题的立场。人们会继续受到自身直觉及情绪的驱动。[注]

海德那篇有名的“情绪化的狗”的文章，开篇用一个例子讲述了推理在影响道德观念时的无力感。

> 朱莉（Julie）和马克（Mark）是姐弟。大学放暑假的时候他们同游法国。一晚，他们独自待在海滩边的小木屋里。于是，他们想，如果他俩做爱的话应该会有趣又好玩。至少，对他俩都会是一种全新的体验。朱莉事先吃了避孕药，不过，为了安全起见，马克还是用了避孕套。他俩都很享受这个过程，但发誓只此一次，下不为例。那晚的事成了他俩的小秘密，为此，他们还感觉彼此更亲近了。对此事你怎么看？他俩做爱合适吗？(注)

海德及其同事与参与者进行了面谈，大多数人并不认可朱莉和马克做爱这件事。当被问及原因时，他们列出了许多理由。平均每个人列出了9条理由，所有理由最终都被实验人员否决了。“他们的小孩可能会有生理缺陷”——不会的，他们用了两种避孕方式，不可能怀孕生小孩；“他们心理上可能会有嫌隙”——恰恰相反，那次体验之后他们反而更亲近了；“若人们发现了，会躲着他们”——他俩都保守秘密，没有人会发现的；等等。但是，参与者没有说，“我用理性解释不清楚自己的道德谴责，因此我收回这一谴责”。大多数人还是坚持自己的观念，即便再也找不到理由来支撑。

不过，互动论方式预料到，理由若合理，应该有些分量。那么，为什么这些人反驳不了海德的论据，却还是固执己见呢？(注)更普遍来说，如果人们对道德论据的回应是不改变自己的看法，那么他们为什么会不情愿呢？

虽然没能改变他人的看法让人既恼火又失望，特别是关于重要的道德问题，但这并不是说不改变看法的人就是不理性的。互动推理理论认为，人们应该会对强有力的理由很敏感，但即便是看似强有力的理由，也不应覆盖其他考量。譬如，我们直觉上可能非常不情愿接受给定的结

论。有些直觉很难弄明白，所以若要我们解释为什么拒绝接受看似有力的论据，我们会一片茫然，不知所措。这并不是说这些直觉就是不理性的，虽然在遇到强有力的论据时没能为自己的立场成功辩论，可能会让我们看起来不理性。某些直觉在阻止我们接受不能有效反驳的论据时，发挥了关键性作用，这些直觉其实与听从专家意见有关。

譬如，若莫阿钠（Moana）试图说服他心胸更宽广的朋友泰基（Teiki），让泰基相信气候变化是个骗局，他俩都会听从专家的意见，但不是有相同背景的专家。听从专家的意见这一举动是理性的。若不如此，对于众多没有遇到过且不能有效思考的重要问题，我们就会毫无头绪。一旦听从某些专家的意见之后，人们不怎么重视的第三方的反驳论据也就说得通了。即便自己可能反驳不了，但我们相信专家可以做到。譬如，莫阿钠的许多论据泰基都不能当场反驳，因为他不太清楚为何自己信任的专家会相信气候变化，但泰基可能不会就此改变看法，因为他觉得自己信任的专家会知道怎么反驳。

若观念不容易验证，基于信任认同它们就是十分理性的，相信不同权威的人坚决反对它们同样是十分理性的。并不是说这些态度就最理性。更机智的做法是让人清楚地看到证据及反对论据，让人心甘情愿地改变看法。随着科学的发展，这一变得更加普遍的做法在认知上就是最好的，但是没人有时间和精力将之应用到每个话题。

论证如何帮助彻底弄清道德问题

我们应该继续推理道德问题吗？孤独推理的影响让人吃不准，论证也面临众多阻碍。但是我们的回答是肯定的。其实，我们觉得比起对乱伦的直觉，大多数道德观念都更经得起论据的检验。为摆平犯罪问题，警察应该怎么做，罗斯（Ross）欺骗雷切尔（Rachel）是多大的错误，这些问题在社群中找不到异口同声的回答；别人同样无法知晓我们是朋

友或是敌人。若持反对意见的人最关心的是把事情彻底搞清楚，论证则不仅能让他们改变看法，还能让他们出于好意改变看法。

检验这一预言时的明显问题在于缺乏明确的道德标尺，来判断论证是否会促成更好的道德观念（根据定义），如果存在明确的道德标尺，那么就应该没有理由辩论了。不过，可以在征得家长同意的情况下，研究儿童案例，探察他们的成长过程发生过什么。如果对某个给定问题，某特定年龄段的孩子有不同意见或表示困惑，便可以看看哪些孩子更有说服力：是那些赞同成人判断的孩子，还是那些持相对欠成熟观点的孩子？

让·皮亚杰先会巧妙地把孩子弄糊涂。譬如，他会给9岁的孩子看下面两个故事。

> 故事一：
>
> 曾经有个叫约翰（John）的小男孩。他待在自己的房间里，后来妈妈叫他吃晚饭。他打开门去客厅吃饭，但在门后有个托盘，托盘上还放着六个杯子。约翰可能并不知道托盘在门后面，开了门，打翻了托盘，六个杯子都打碎了。
>
> 故事二：
>
> 曾经有个叫大卫（David）的小男孩。一天，妈妈出去了，他试图从食品柜里拿一些甜食，爬到椅子上，伸出手臂。但是甜食放得太高了，他够不到，在试图够到（甜食）时打翻了一个杯子，杯子落下来摔碎了。㊟

皮亚杰会询问孩子：这两个男孩，约翰和大卫，哪一个更淘气呢？当帕特里克·勒曼（Patrick Leman）和杰拉德·杜维恩（Gerard Duveen）重做皮亚杰的实验时，他们发现，大多数9岁的孩子都认为约翰更淘气些。㊟作为大人，我们现在或许都会认为孩子们给出的答案是错误的。

大卫到底有多淘气是一个待讨论的问题，但显然约翰什么也没做错。他打碎杯子纯粹是巧合，连粗心大意都算不上。令人欣慰的是，当实验人员让持不同看法的两个孩子讨论时，他们最终认为大卫更淘气些的可能性高达五倍。多亏了论证，道德判断才变得更准确。

民众政治辩论的神奇功效

成人之间在讨论时，某些道德性辩论通常会变成政治性辩论，不仅辩论是非对错，还辩论社会解决这一问题的途径。人们很容易对政治辩论形成模糊的观点。在某些民主国家，最公开的政治辩论通常发生在总统竞选人之间。这些辩论是有些不自然的表演，辩论双方都知道不可能说服对方，主要只是加强对自身立场的支撑而已。幸好，对政治问题的辩论不只发生在总统竞选人之间，市民之间也会进行辩论。

萨缪尔·亨廷顿（Samuel Huntington）主张“开放、自由、公正的选举是民主的实质”时表达了一种普遍性观念。[注]但是，投票不是公正地汇集观点的唯一方式。真的，它既不是最古老的形式，也不是最普遍的形式。[注]正如我们在第十六章提到的，狩猎采集者群体基于公共协商做出群体决定。也就是说，这些群体里的人与我们旧石器时代的祖先有某种相似性，这就是说协商比投票更久远。阿玛蒂亚·森（Amartya Sen）在其《民主及其全球根源》（*Democracy and Its Global Roots*）一书中，让读者简要了解了非西方民主惯例，其中许多都是协商性质的。[注]印度君主阿克巴（Akbar）16世纪发起的巨大的跨宗教性辩论，给青年时期的纳尔逊·曼德拉（Nelson Mandela）留下了“最纯粹民主”印象的腾布王朝开放会议，[注]遍及全世界的协商承载着形成更好观念、做出更好决定的希望。

20世纪80年代初期，政治学家开始更多地关注协商在健全民主中发

挥的作用。[注]起初，协商民主（deliberative democracy）这一新领域关注带有“高尚理想”的方面，关注协商在促进理性话语、公民性、公众参与度、相互尊重等方面的潜力。接着，政治学家让这些“高尚理想”直面现实中的协商，这些协商是在有分歧、被误导，有时甚至极愤怒的民众之间展开。出乎众人的意料，崇高的理想获胜了。当把一些民众聚在一起、分成小组，在仲裁人的友情点拨下一起讨论政策问题时，好事便发生了。[注]参与者经过讨论，最终消息更灵通、立场更鲜明，对他人立场也了解得更深入。他们的观念汇聚在一起，最终形成了合理的折中观念。他们未来参与公众生活的可能性更大。市民间的协商起了作用。

其中一项最成功的协商民主实验是由罗伯特·拉斯金（Robert Luskin）和詹姆斯·费希金（James Fishkin）发起的。他们在众多城市进行了协商民调（deliberative polls），让市民一起讨论，最后这些市民能就各种政策问题形成更理智的立场。其中一个城市便是北爱尔兰的奥马。

1998年8月15日，奥马发生了炸弹袭击，29人遇害，200多人负伤。据称袭击者是爱尔兰共和军（Irish Republican Army）分裂出来的小派别，被极富创意地命名为真爱尔兰共和军（Real Irish Republican Army），这次袭击是北爱尔兰统治下长时间流血冲突中最惨烈的一次暴行。在奥马，天主教徒和新教徒有充足的理由怀疑彼此、坚持本教信仰，这不是协商发挥作用的最好地方。

然而，当拉斯金、费希金和另外两个同事一起，让包括天主教徒及新教徒在内的当地人一起讨论教育政策时，辩论很有建设性意义，即便非常沉重的话题亦是如此。[注]当混合型教会学校问题出现在辩论中时，参与者也不会斗争、做出极端之举。讨论过后，参与者不仅就几个问题改变了看法，还更深入了解了教育政策。他们也发现对话者比想象中更值得信任，更开明。

批评家指出协商民主正扩大化的问题是：在少数人中辩论可以开展得很好，但在好几百万人中就没那么好了。费希金与美国宪法学者布鲁斯·阿克曼（Bruce Ackerman）一起，正式提议设定协商日，即法定节假日，在这一天邀请市民对即将开展的选举进行辩论。虽然这种制度会进一步增强讨论在公众生活中的作用，论证其实已经证实了其影响大范围道德及政治性变革的能力。

废奴主义：做出论断并不容易

18世纪末，英国人主导了横跨大西洋的奴隶贸易，(注)还在美洲获得了大块领土，带着获取无尽财富的期许。受经济利益的驱使，他们应该俘获并运送成千上万的奴隶来开发这些土地。(注)但相反的是，他们选择废除奴隶贸易。废奴主义者是怎样实现这样一个大反转的呢？

我们现在认为，这似乎是一个很容易做出的论断。为什么说服人们相信奴隶制度大错特错、应该被废除会很有必要呢？可惜，奴隶制度的恶行并不总是道德上的自明之理。真的，在大部分历史中，奴隶制度已成为人们生活的一部分。希腊人及罗马人亲身实践，犹太教、基督教及伊斯兰教大力支持，欧洲历史的大部分时段都对奴隶制度习以为常，不认为这是一个问题。当狄德罗（Diderot）等启蒙运动的先驱坚决地谴责这一行径时，这一历史大潮改变了方向。然而，新宗教运动，尤其是贵格会宗教运动，为《圣经》（*Bible*）提供了新的解读，把奴隶制度变成一种非常野蛮的制度。奴隶贩子及奴隶主终究要给出支撑自身行径的正当理由。

奴隶制度的辩护者前来帮忙，甚至一度试图占据道德制高点。他们辩论称，非洲的生活太艰苦了，让人简直无法忍受。相对而言，在去往美洲的“中间航程”（Middle Passage）里，奴隶被当作贵宾，享有“香酒……烟斗及烟草”，并“欢快地听着各种乐器声”。奴隶贩子声称，“十

分之九的（奴隶）很乐意落到我们手中”。[注]再者，整个奴隶贸易只是各取所需，因为奴隶一旦抵达目的地，其血脉便难以维系。那是因为女性奴隶成了“妓女”，必须要频繁“堕胎才能继续交易，不至于浪费时间”。爱德华·朗（Edward Long）在其《女性史》（*History of Famaica*）一书中继续说道，“这种滥交必定会降低怀孕概率或让女性怀不了孩子”。[注]奴隶应该对奴隶贸易心存感激，但也要为此负责。

这些论据听起来不仅糟糕，而且荒唐。可是那时可靠信息匮乏，英国公民不了解在非洲、西印度群岛或美洲发生的事。而且，在英国的生活恰好也不太容易。工业革命虽然创造了巨大的财富，但也造成了极大的不幸。在英国各个港口，成千上万的人被“强征”、绑架，然后带上海军军舰，遭受“好几年的鞭笞，染上坏血病和疟疾”。[注]由于奴隶贩子描绘的境况比较美好，所以一般人绝不会想到奴隶的境况竟如此糟糕。

但是，反对废奴的人拿出的最有说服力的论据也不是道德上的，而是经济上的。像利物浦这类城市的整个运转都依靠奴隶贸易。即便像曼彻斯特这类内陆城市，也需要依靠奴隶在殖民地收集原材料，有了源源不断的原材料供应，才可以雇佣纺织工人。奴隶贩子的代言人认为废奴之路会让整个大不列颠满是“孤儿寡母”，并不厌其烦地强调“孤儿寡母”这一信息。[注]反对废奴的人信手捏造一个数字，称此举会造成7 000万英镑的损失！[注]他们还会拿英国人最爱喝的酒说事：哪能没有“糖和朗姆酒”，“简直不能忍受茶的味道”。

然而18世纪80年代中期，贵格会信徒及其他早期的废奴主义者已经成功夺回了道德制高点。他们用了某必不可少的论证手段，即显示受众立场的不一致性。在这种情况下，这种不一致性十分明显：一方面表现在基督教和英国精神中，另一方面表现在奴隶制度上。詹姆斯·多雷（James Dore）于1788年布道时曾辩称：“作为人类，胸中应该激荡起对奴隶主之间进行奴隶买卖的憎恶，特别是作为大不列颠人，尤为重要的是作为基督徒，应该憎恶这种行径。”[注]历史学家西摩·德雷舍

（Seymour　Drescher）指出，这种不一致性已成为盛行的废奴主义的主要推动力量：“作为世界上最安全、最自由、最虔诚、最公正、最繁荣且最道德的民族，怎能让自己同时也是世界上最致命、最残酷、最不公正、最不道德地冒犯人性的主犯呢？”[注]尽管如此，奴隶贩子提出的经济考量还是岿然不动。道德论据太抽象了，虽然奴隶制度给奴隶带来了巨大且深远的苦难，但要理解这种苦难还是比较费劲的。废奴主义者需要找到更多证据，来让其论据充分发挥作用。

说服一个国家

几年来，废奴主义者托马斯·克拉克森（Thomas　Clarkson）多次往返英格兰，收集奴隶贸易留下的众多证据。其成果是：1790—1791年将证据摘要呈递给下议院的特别委员会看，这一摘要是由废除奴隶贸易的请愿者完成的。这些成果便成为废奴主义者日益壮大的军械库里的主要武器。这一军械库的完工离不开创造“敏感性修辞”的男男女女，他们创作旨在恢复奴隶完整人性的诗集；[注]出力的还有解放了的奴隶，他们写的自传很成功并广为流传，他们大力宣传证据摘要里的数字；昔日的奴隶贩子也有一份功劳，他们见证并佐证了奴隶生活的惨状，让奴隶贸易的惨状具有公信力。不过，证据摘要仍然是“英国全民动员的核心文献”。[注]但是，除了议会高墙外大众的诉求，废奴主义者还需要听到圈内人的意见。

威廉·威尔伯福斯（William　Wilberforce）虽为议会成员，但对许多问题都持保守观念。18世纪80年代中期，他变成了福音派信徒，这一信仰上的转变似乎让他开始更积极地回应废奴主义者给出的论据。这场运动也游说了威尔伯福斯，最终说服了他，让他为此发声。与其他废奴主义者一样，威尔伯福斯指出，奴隶制度和民族归属感之间存在不一致性，这个民族“除了无与伦比的、真正的公民自由，还沐浴着空前的宗

教光辉，承载着教众满满的祝福”。[注]威尔伯福斯不只演练了标准性论据，还掌握了证据，熟悉了反对废奴者的论据，并采用他们的立场来与之辩论。奴隶贩子声称，虐待自己最宝贵的货物毫无经济价值。威尔伯福斯指出，与奴隶贩子所言相反，“商人要获利，奴隶的数量很关键，而且还只能给奴隶少量的补贴”。[注]朗在《女性史》中的详尽阐述大概可以为反对废奴的人提供许多可靠的论据，威尔伯福斯就决定用朗的断言来作为自己的论述的前提。朗的断言是，“黑人用最少或最简单的劳动培育了最好的物种”。[注]威尔伯福斯补充道：“好吧，要是奴隶主稍稍仁慈一点就好了，这样，奴隶的人数便不会骤减，就没有必要进行奴隶贸易了。”

废奴主义者收集的众多理由及证据最终说服了大多数议员——直接地或通过论据所搜集的民意。1792年，四分之三的下议院议员投票赞成逐渐废除奴隶贸易。而上议院议员则与奴隶贩子的利益更贴近，称还要仔细考虑这个问题。其实是时机不合适：多年来，法国革命的风险、拿破仑（Napoleon）的强势，导致所有激进运动被镇压，那时就包括废除奴隶制度的倡议。但一旦有机会，废奴主义者便会完善其论据，重新唤起大众的诉求，让议会充斥着新的请愿书。而且，威尔伯福斯再次轻而易举地掌控了下议院的辩论走向。1807年，上议院和下议院都投票赞成废除奴隶贸易。

英国的废奴主义者提出的大多数反对奴隶制度的论据都并非杜撰，只是对这些论据加以改善，用大量的证据来予以支撑，借助可靠的目击者来增加这些论据的可信度，为了让这些论据更容易理解，废奴主义者还让人站在奴隶的视角来看待生活。辩论、公众集会和报纸让这些经过改善的论据呈现在蓬勃发展的城市人口面前。而且，这真的有用。人们不仅确信奴隶制度的危害性，还认为有必要做些什么来改变这一现状。人们请愿、捐钱，并借助其他因素（无论是经济因素还是国际政治因素）先让奴隶贸易被禁止，而后让奴隶制度被废除。

理性最好和最坏的一面

互动理论站在独特的视角，来解释理性对道德观念及决定产生的影响。许多实验及先前无数个人的历史性观察都让唯理智理论的道德理性观念站不住脚。道德观念及决定通常受到直觉和情绪的支配，理性至多只会提供滞后的合理化，最坏也只是给推理者做出道德上可疑行为提供借口，如不坐在身有残疾的人旁边、保留奴隶制度。作为有偏见且懒惰的正当理由制造者，理性只是做了人们预料之中的事。

但是，关于理性改变人们看法的能力，我们不太认同悲观主义的观念。人们不仅给出自己的正当理由及论据，还评估他人给出的正当理由及论据。作为评估者，人们不仅要能辨识强有力的论据，而且要在任何领域都能受其影响，包括道德领域。显然，若某个论据挑战社群的道德观念，则只会让人怀疑其动机，得不到信任，甚至还会让人产生十足的敌意。尽管如此，人们在许多道德问题上还是受到合理论据的影响，如怎样组织地方学校课程等地方政治性问题、废除奴隶贸易等重要的社会性问题。

---

1. Schnall et al. 2008.
2. Danziger, Levav & Avnaim-Pesso 2011;但是要了解对这些结果的怀疑性解释，可参见：Glöckner 2016。
3. Haidt 2001.
4. Snyder et al. 1979.
5. Chance & Norton 2008.
6. Rust & Schwitzgebel 2013.
7. 也有许多其他的例子，可参见：Uhlmann et al. 2009; Valdesolo & DeSteno 2008。
8. “产权”问题可查看Jefferson Monticello网站，网址：http://www. monticello. org/site/plantation-and-slavery/property。
9. Thomas Jefferson to John Wayles Eppes, 1820.6.30，Founders Early Access, 网址：http://rotunda.upress.virginia.edu/founders/default.xqy?keys=FOEA-print-04-02-02-1352。

10. Stanton 1993, pp.158-159.

11. Letter to Thomas Mann Randolph, 1803.6.8, in Betts 1953, p.19.这一特别的引用只指代一个奴隶。

12. 看他们如何扭曲历史记载，可参见：Finkelman 1993。

13. 所有的引用都来自《第十四次质疑》（*Query XIV*）, 可以从许多在线来源中找到；可查看网址：http://xroads.virginia.edu/~hyper/jefferson/ch14.html。

14. 所有引用都来自《第十四次质疑》（包括“飘逸长发”这个引用）。

15. 他的这一立场既促使他没有解放即便是一小部分奴隶，也向其他奴隶主提出了建议，让他们也不要解放奴隶。至少这是托马斯·杰斐逊于1814年8月25日写给爱德华·科尔斯（Edward Coles）的信中给出的建议；可参见：American History: From Revolution to Reconstruction and Beyond，网址：http://www. let.rug.nl/usa/presidents/thomas-jefferson/letters-of-thomas-jefferson/jefl232.php。

16. 他的这一立场既促使他没有解放即便是一小部分奴隶，也向其他奴隶主提出了建议，让他们也不要解放奴隶。至少这是托马斯·杰斐逊于1814年8月25日写给爱德华·科尔斯（Edward Coles）的信中给出的建议；可参见：American History: From Revolution to Reconstruction and Beyond，网址：http://www. let.rug.nl/usa/presidents/thomas-jefferson/letters-of-thomas-jefferson/jefl232.php；也可参见：Finkelman 1993。

17. 出自《第十四次质疑》。

18. Letter from Thomas Jefferson to Edward Coles, 1814.8.25.

19. Finkelman 1993, p.186。更为普遍地说，种族主义为奴隶贸易辩护，在发展时遇到了人人平等理想的阻扰，欲了解有关信息，可参见：Davis 1999; Fields 1990。

20. Franklin 1799;着重强调。

21. 杰斐逊的情况也说明了，就算意识到了偏见，也不会减少偏见。正如他所言，“一旦一个人的理论形成，他就只能从每个物件中找到支撑该理论的特性”。（Jefferson 1829, p.235）

22. Haidt 2001, p.823.

23. 但是可参见：Haidt & Bjorklund 2007。

24. Haidt 2001, p.814；可参见：Haidt, Bjorklund & Murphy 2000。

25. 关于“道德错愕效应”的局限性，可参见：Royzman, Kim & Leeman 2015。

26. 来自Piaget 1932, 改编自Leman & Duveen 1999, p.575。

27. Leman & Duveen 1999.

28. Huntington 1993, p.9.

29. 除非人们用自己的脚投票；可参见：Conradt & Roper 2005。
30. Sen 2003.
31. Mandela 1994, p.21.
32. 可参见：Bessette 1980; Cohen 1989。
33. 可参见：Fishkin 2009; 还可参见：Landemore 2013; Mercier & Landemore 2012。
34. Luskin et al. 2014.
35. Drescher 2009, p.207.
36. Drescher 2009，p.224。
37. 语出理查德·米尔斯（Richard Miles），引自Sanderson 1972, pp.68, 71；也可参见：Hochschild 2006。
38. Long 1774；也可参见：Perry 2012。
39. Hochschild 2006, p.223.
40. Perry 2012, p.95.
41. Perry 2012，p.90。
42. 引自Swaminathan 2016。
43. Drescher 2009, p.213.
44. 可参见：Carey 2005, pp.85ff。
45. Drescher 1990, p.566.
46. Wilberforce 1789.
47. Wilberforce 1789，p.52, 引自Perry 2012。
48. Long 1744, p.437，引自Perry 2012。

第十八章
孤独的天才？

20世纪20年代，一些粒子物理学家在研究时遇到了进退两难的困境，一着不慎，后果会很严重。他们借助数学运算来理解基本粒子，但这一数学运算与基本粒子的标准空间表征相冲突。虽然数学运算有助于做出关于粒子行为的精准预测，但这些等式还不能算出粒子的空间运行轨道。量子论似乎存在很大瑕疵。物理学家要否决这个相当成功的架构吗？还是要被迫接受其局限性呢？沃纳·海森堡（Werner Heisenberg）已成为这一群体中知名度很高的一个人物。他创造了解读基本粒子的数学架构，对产生这一窘境形成推波助澜之力。所以，海森堡着手解决此事。

那时，对粒子物理中大多数运动的研究都是在尼尔斯·波尔（Niels Bohr）位于哥本哈根的实验室中进行的。波尔不仅是一个伟大的科学家，还具备筹集资金的天分，他妥善地利用这笔钱给海森堡等年轻有为的研究员增加研究经费。然而，波尔具有传奇色彩的人设与海森堡的雄心和独立情怀冲突了。为寻求和谐，海森堡将研究阵地转移到戈兰高地某小岛上，距德国海岸约三个小时的航程。独立研究数月后，海森堡终于找到了应对量子物理常遇到的这一窘境的解决方案：一个表示不确定性原理的数学公式。海森堡用外行人的说法将这一原理概括为："有两个重要因素决定了其中一个最小粒子的运动（其位置及速度），而人们永远无法同时十分确切地了解它们。"[注]这是第三条道路：不必用新架构取代量子物理，物理学家也不必屈从于不完美的世界观。如果不能推算出基本粒子的精确运行轨道，不是因为量子物理存在瑕疵，而是因为

位置和速度在测量前并没有同时存在，并且不能被同时测量。

在小岛上与世隔绝后，海森堡的卓绝才智才得以施展出来，进而更深入地理解世界的本质，完美阐释了孤独天才的科学观。从这一通俗观念来看，才智卓绝之士在隔绝的环境下工作，往往能产生科学性突破。孤独天才的形象是18世纪浪漫主义虚构出来的，可以在诗人华兹华斯（Wordsworth）的诗歌中找到渊源。对他而言，牛顿的雕像是：

> 思想永远的大理石索引
>
> 独自一人，在陌生的思想之海上航行[注]

在孤独天才的科学观里，学生、助手和不那么优秀的科学家都是走卒，这些人是用来提供数据的，天才十分认真地思考问题，由此形成漂亮的理论。这与弗朗西斯·培根提倡的科学观非常相近。

培根是个有远见的政治家及哲学家，16、17世纪之交时生活在英格兰。他力劝同事摒弃华而不实的论据，采用更务实的方式，让他们不要再对亚里士多德的观念妄加评论，要开展实验，这些想法会激发科学革命。虽然培根强调了科学的合作性特点，但他眼中的合作是分等级的。在描述空想的大西岛（New Atlantis）时，培根列出了一长串名单，让这些人负责“收集书本中所有的实验……（以及）机械艺术中所有的实验”，并让这些人“尝试新实验”。但理论化工作就留给其精心挑选的小组成员：“最后，我们要让先前实验中的三个发现变成更伟大的观察结果、原理和警句。我们将之称为‘天性解读者’。”

培根生活的年代有两个最优秀的律师，而他则是其中之一，所以通晓论证，不过，其宏伟计划中却没给论证留下多少余地。若无论怎样都要论证，存在等级之分的科学观便使公认的天才能够说服比其能力弱的人相信其发现。即便效果可能没有那么好。

由于大众接受其观念比较缓慢，马克斯·普朗克（Max Planck）感到气馁，作为量子物理学的奠基人之一，他讥讽道："新的科学真理得到认可不是因为说服了其对手、让他们认可并领悟，而是因为最终没有了对手，而且新一代人在成长过程中对该真理日渐熟悉。"[注]普朗克的警句概述了其对科学变革（或科学变革的缺乏）根深蒂固的信念，在成千上万的网页、书籍及科研论文中得到引用。

若科技进步的推手是建造革命理论的孤独天才，即便这些天才也没能说服同事认同其构想，那么，科学，我们的认知探索中最成功之物，便违背了我们互动推理理论的所有预测。那将会是一个问题。

科学家也是有偏见的

科学家会借助各种认知技能。正如庞加莱（Poincaré）、爱因斯坦等研究员指出的那样，要诞生新的真知灼见，直觉的力量不容小觑。当然，没有人否认理性对科学的重要贡献。科学家独自便能回答十分复杂的问题吗？如果能做到，他们一定是十分优秀的推理者，能克服困扰外行人的两个问题：一是自我中心偏见，二是以较低的标准评估自己的论据——懒惰。或许科学家比较特别。或许他们对证伪已经再熟悉不过了，几乎都快成了第二天性，让其在对自身观念进行推理时能够秉持公正的态度。第十一章中讨论过莱纳斯·鲍林的案例，该案例让我们得以瞥见答案，但是要证明该问题，同样的逸事证据还比较匮乏。

20世纪70年代，迈克尔·马奥尼（Michael Mahoney）及其同事开展了一系列的实验，就为了探求科学家是否也有自我中心偏见。他们在其中一项研究中设计了一项评估自我中心偏见的任务，把科学家与新教牧师在心理学及物理学这两个不同领域的回答进行比较。三个小组都表现出了偏见，与牧师及先前实验中的参与者一样，科学家也有偏见。[注]在马奥尼开展的另一项实验中，他让科学家对自身领域的专业技能进行推

理，这次，他又察觉到了强烈的自我中心偏见。㊟

科学家在自然环境中的观察结果也出现了同样的情况。凯文·邓巴（Kevin Dunbar）特别研究了科学家在现实生活中的思维模式，而不只是在实验环境下。㊟为了弄清楚研究员处理数据、提出新假设及解决新问题的方式，他与许多研究员进行了面谈。其观察结果表明，科学家会运用推理来摒弃比较麻烦的结果。当实验的结果不尽如人意时，研究员不会进行公正的推理，不会质疑自己最初的假设，也不会提出新的假设，而是满足于以下三条微弱无力的论据，幻想用它们来捍卫最初的假设：其一，出现了技术性问题；其二，实验有瑕疵；其三，有人出了差错。

科学家也对合理论据敏感

若与外行人的推理一样，科学家的推理也有偏见和局限性，那么也应该具备外行人推理的优势，即与制造理由相比，评估理由时要客观得多。科学家应该关注彼此的论据，若他人所给理由比较合理，就适时改变自身看法。但科学家似乎连效仿普朗克、遵循常识都做不到，就算只剩最后一口气，也要坚决抵抗新理论。

平心而论，科学家确实被迫接受了一些疯狂的想法，如我们是单细胞生物的后代，是数十亿年的随机突变的产物，能绕着太阳以每小时十万多千米的速度快速移动；使我们黏附在地面的力量与让地球不偏离自身轨道的力量相同；多亏光束粒子，我们才能看到眼前这页论述，时间因它而不复存在，我们的意识是三磅脑灰质的产物。普朗克努力说服物理学家相信能量不是连续性变量，而是以指定大小的量子形式出现，与此同时，他也在抱怨自己的观念扩散得太慢。要创造出量子物理理论，普朗克就要违背经典物理学中一项最基本的假设。所以有人怀疑也是无可厚非的。科学观念违反常理不是减慢其传播速度的唯一因素。革命性

书籍《科学革命的结构》（*The Structure of Scientific Revolutions*）的作者托马斯·库恩（Thomas Kuhn）认为，抗拒新理论的大多是把全部心血投入一项研究的科学家。对于新贵明显荒唐且危险的想法，科学家虽有怀疑，但不会因此指责新贵。

虽然与先前所有人相比，库恩更加意识到让革命性观念得到认可所面临的挑战，但他并不认同普朗克单薄的评估：“尽管一些科学家，特别是更老更有资历的科学家，可能会一直抗拒下去，大多数科学家不管怎样还是会受到影响的。”[注]讽刺的是，另一位支持科学革命的学者伯纳德·科恩（Bernard Cohen）指出，“普朗克本人实际上见证了自身基本概念被同事认可、修改及应用”的方式。[注]普朗克和库恩认为年龄会对人们从容接纳新观念的能力产生不利影响，这一观念似乎也是错误的。历史性定量研究表明，与年轻的同行相比，年长的科学家更有可能去接受新理论。[注]

科恩评论科学革命时进行了进一步阐释。他指出，在科学革命中，不仅观念得到了转变，而且，“无论是谁，只要读到科学革命文献，随处可以见到信仰的转换”。[注]遇到合理的论据，有时会让我们一下子顿悟真理，正如19世纪某化学家描述的那样（该化学家在某次会议上获赠一本小册子），他说，“我……在家里反复地看这本小册子，惊奇地发现这个小东西竟然阐明了最重要的争论点。随着谜题渐渐揭晓，疑虑消失得无影无踪，心里无比平静，十分确定”。[注]

科学革命以温和的方式继续进行。没有必要砍掉资深科学家的头；他们可以“心甘情愿地认同”[注]，然后改变自身观念。达尔文发表《物种起源》（*The Origin of Species*）后过了10年，四分之三的英国科学家至少在某种程度上已经被说服了。[注]离有说服力的板块构造理论得到一致认同已经过了10多年，该理论早已编入了教材。[注]

科学充分利用论证

科学家的推理与外行人的推理是同一种类型的推理。科学进步不是由特殊的超级推理者群体来实现的，而是通过充分利用推理的优势，即促进讨论、为人们提供辩论工具、让人们有选择性地改变自身看法。

皇家学会（The Royal Society）是世界上第一批科学学会之一，于1660年在英格兰建立。其好几个创立者都成了科学革命中的领军人物。然而，不论他们多么才华横溢，还是在某些事情上出了错。譬如，克里斯托弗·雷恩（Christopher Wren）相信“好探究的哲学家会发现真正的占星术”；罗伯特·波义耳（Robert Boyle）认为，刚被绞死的人的手是治愈甲状腺肿的药剂；约翰·威尔金斯（John Wilkins）认为瓦伦丁·格瑞克斯（Valentine Greatrakes）具有神奇的力量，是“17世纪最著名的神秘治疗师”[注]，可以治愈肺结核及其他疾患。但是，他们正确的信念比其任何特定的错误信念更加重要。这是人们获得全新且更好的信念的途径。这些先驱认为，真正的实验应该代替观念性实验[注]，认为学者应该进行开放性对话，而不是投入像中世纪辩论一样缺乏新意的智力游戏中，他们还认为，与沿用几百年的书籍相比，从商人及旅客那儿能获取更多的信息。

雷恩、波义耳、威尔金斯及其同事目标定得很高。为阐释讨论的力量，波义耳写就了《怀疑派化学家》（*The Sceptical Chymist*）一书，虚构了化学家之间的研讨会。对话风格已经屡见不鲜了，因为伽利略（Galileo）曾大力运用这一风格来阐发他自己的革命性理论。但是在其《两大世界体系的对话》（*Dialogue Concerning the Two Chief World Systems*）一书中，很明显，伽利略的代言人萨尔维亚蒂（Salviati）一开始就把一切都弄清楚了。

波义耳所用对话的形式有所不同。在《怀疑派化学家》一书中，没有一个人物一开始就知道答案。这只是“一出戏剧，展现出劝说、意见

分歧及最终向真理的转变是如何实现的”。㊟主角卡尼阿德斯（Carneades）的观点虽是波义耳观点最贴切的反映，他却没有把这一事实灌输给对话者，“而是让这一事实在戏剧中表现为是通过对话产生的”。㊟

波义耳宏大的构想可能并没有付诸实践，但是，对当代科学每天的活动，他还没有感到彻底失望。实验已成为必然趋势。而且在每天成千上万的实验室会议中，人们也能察觉到波义耳有教育意义的对话。邓巴发现，当产生不尽如人意的结果时，科学家会当机立断为之辩护，他们到底遇到了什么？当科学家在实验室会议中给出这些托词时，遭到了断然的拒绝。合理化的合理性只够说服做出这一合理化的人，还不能说服其他实验室成员。鉴于其他小组成员提供了选择性假设及解释，研究员只得在这些人的协助下，对实验结果做出更好的解释。㊟

邓巴在实验室会议中观察到，讨论约束误导个体推理者的偏见的方式得到了完美展示。而这只是科学范畴内论据交流所采取的众多形式之一，这些形式就包括非正式交谈、同行评审及国际研讨会。

当代科学家一直在论述讨论及论据对科学的重要性，这一观念与诗人及其他外部观察者不同，这些人创造了孤独天才的神话。某数学物理学家接受了米哈里·希斯赞特米哈伊（Mihaly Csikszentmihalyi）和基思·索耶（Keith Sawyer）关于创造力问题的采访，将科学描述为：

> 它非常具有社交性，本质上展现了打开这扇门与关上这扇门的差异。如果我正在搞科研，便把门打开了。这有点象征性，却是真实的。你想要一直与人交谈……只有与楼里的其他人互动，才能做出有意义的事情。它本质上是一项公共事业。㊟

丹尼尔·卡尼曼与阿莫斯·特沃斯基对促进个人主义推理观念的发展做出了举足轻重的贡献，但即便是丹尼尔·卡尼曼，也充分意识到了讨

论的力量，他说："我在与阿莫斯悠闲地散步时，对自己的人生思考得最透彻。"[注]

孤独的天才有多孤独

科学社会学家与科学历史学家一致认为，科学本质上是一项集体性事业。不论是培根眼中卑微的数据收集者与傲慢的理论家之间的劳动分工，还是协同性更高的尝试，科学都掺杂大量的讨论及论据。不过，似乎存在例外情况。你怎么看待那些独自做出新突破的科学家？又怎么看待在小岛上做研究的海森堡？

科学历史学家玛拉·贝勒（Mara Beller）认真观察了让海森堡形成不确定性原理的过程，从而拒绝接受孤独天才这一观念。贝勒通过参阅科技论文、读者投书栏、采访和自传等文章，重构了海森堡成果的对话性。贝勒认为："海森堡之所以对物理学做出了不朽贡献，不是因为他权威性地延伸了单一论据，而是因为创造性地合并了不同论据，这些论据彼此强化、相互阐明。"[注]海森堡的深刻见解融入了波尔（Bohr）和狄拉克（Dirac）的思想，是对薛定谔（Schrödinger）立场的回应，还借鉴诺曼·坎贝尔（Norman Campbell）和H.A.森特福利本（H. A. Sentfleben）等没什么名气的物理学家的观念。[注]

最重要的是，海森堡与沃夫冈·鲍理（Wolfgang Pauli）的观念高度一致。在正视了鲍理的论据并应对他的挑战后，海森堡便能使自己的观念充分发挥潜力。海森堡关于不确定性原理的论文初稿的确在他给鲍理的一封信中出现了。即便是天才，也需要与人辩论，这样才能形成最好的观念。

海森堡与鲍理之间的论据交流对不确定性原理的构想极其重要，但这种论据交流远非日常对话。我们每天通过独立思考，就会找到一系列

组织严密的论据，而非短小的、来回都是那么几句的陈述。由此可以得出，创造科学性突破的思维过程与心理学实验描绘的单独推理不同，它没有无限堆积以支持科学家观念的未经检验的论据。自我中心偏见的确存在，但是这一偏见被更严格的质量控制打磨了，最没有说服力的论据被从中剔除。

我们在第十二章中主张，推理者在日常对话中不花费太多精力预测可能的反对论据也是说得通的。既然没有立即说服对话方所付出的代价不大，那么预测反对论据便变得既困难又用处不大。但是，在整个进化过程中，推理不可能只在这些情况下得以运用。有时风险要更高些，在某些情况下，反对论据预测起来会更容易。所以，对于人们查看自身论据的数量及细致度，我们应该允许一些变数。

某些机构会重视精心设计的论据。律师只是去做最后的辩护。政客必须要将自己的论据变成合理有效的互撕工具。科学家互相竞争，都希望得到关注，但只有拿出最合理的论据，才有机会得到关注。推理在某种程度上应对了这些挑战。专业人员通过训练已经可以对自己的论据提出相对较高的标准。大多数情况下，这一改善显然是为了坚定信念。律师必须说服法官或陪审团。若律师所给的论据没有说服他人，甚至连自己都说服不了，那也只能这样了。

与之相比，科学家——往好处说——似乎是在为真理而奋斗，而不是为了得到某受众的认可。我们认为，科学家的推理事实上还是在寻找旨在说服他人的论据，可是目标受众群太庞大了，而且要求还特别高。无论是谁，只要能拿出令整个科学界信服的反对论据，科学家给出的论据便不攻自破了。不能投法官所好，也不能有意利用陪审团不懂法的事实。科学家提出的论据必须具有普适性，要能在其所属领域立足。

至少，与谋求普遍认同的动机一样重要的是其实现方式。第十二章中，当伊莲娜试图说服马乔里在伊佐美用餐时，她几乎预测不到马乔里

的反对论据，因为这些反对论据是基于比较特殊的信息，如马乔里上周的用餐地点或她对吃饭花销的承受能力。相比之下，某特定科学领域的专家很可能会共享许多与解决分歧相关的信息。这就让实际论证变得更加有效。这也增强了个人推理的力量。若某科学家在评估自身论据时发现了反对论据，那么其他人也可能发现了；若他没有发现任何反对论据，那也就意味着其他人也不会发现任何反对论据，就算真的找到了反对论据，也不会让相关群体信服。说话人与受众的观念重叠越多，个人推理会越有用。从牛顿到佩雷尔曼（Perelman），这些数学家则处在频谱的末端，比较与众不同的数学家，他们在完全隔离的环境下取得了举世瞩目的新成就。

在所有从事科学研究的人中，数学家最有可能会承认同样的事实，也最有可能会被同样的论据说服，他们用的是同样的定理，定理中的成分也相同，想要给出的论据就是证据。从形式化视角来看，证据是形式化派生物，其中，结论必然是由前提推导而来的。从科学社会学的观点来看，证据就是论据，科学界认为这是一劳永逸的办法。在逻辑和数学中，形式化的与社会学的证据观念往往有同样的范围。[注]1930年，库尔特·哥德尔（Kurt Gödel）推出了其第一个不完备定理，摧毁了为数学建造一致且完整的一套定理的梦想，粉碎了数学家曾设计的最有雄心的计划。它是一个真正带有革命性及威胁性的结果。然而，每个看了哥德尔所给证据的数学家都会很快接受这一结果。[注]

若数学家很快就能达成共识，一致认同游戏规则，那么应该也能非常有效地预测彼此的反对论据。比其他任何小组更胜一筹的是，数学家既有动力，也能详尽地评估自己的论据。群智必然会发现被个人独立思考所忽视的缺陷，但是孤独的数学家仍然有机会取得伟大的成果。佩雷尔曼解决庞加莱（Poincaré）猜想就是在通常无人且破旧的俄罗斯研究所及其母亲的公寓里。[注]安德鲁·怀尔斯（Andrew Wiles）在近乎完全密封的环境中用了6年时间证明费马（Fermat）的猜想。

对话仍能给数学家带来巨大的利益，20世纪就有许多伟大的数学家通过成百上千次合作来达到目的，其中就包括保罗·厄多斯（Paul Erdos）。但是其必要性不如其他学科，即便在其他所谓硬科学中，研究员也几乎不会像数学家那样苛求论据，他们会让自己的论据在讨论中完善。

科学的社会语境推动改善孤独的推理

因为科学为其实践者提供了动力，还让实践者能进行卓有成效的单独推理，所以那些最杰出的科学家似乎具备超自然的能力，得以形成真知灼见，在这方面，牛顿的成就无人能及。著名的牛顿传记作者、历史学家理查德·韦斯特福尔（Richard Westfall）告诉我们：

> 但是，我从未遇见让我不愿意把自己与其进行比较的人，所以我说自己的能力只及牛顿的一半似乎合情合理，或者说只及三分之一或四分之一，但是在任何情况下都是占了一定的比例的。我研究牛顿得出的最终结果让我相信，没有任何衡量标准适用于他。在我看来，他已经完全变成了另一个人，属于为数不多的拓展人类智力范畴的绝顶聪明之才，他没有去迎合我们为一般人设立的标准。[注]

早在韦斯特福尔给牛顿写传记之前，人们就已经把牛顿奉若神明了，[注]但是也有一些早期的怀疑论者，如约瑟夫·普利斯特利（Joseph Priestley），就小心翼翼地将牛顿的成就纳入人类智力成就的范畴：“要是我们能洞悉牛顿的思想，要是能了解他伟大成就产生的每个步骤，在这一过程中我们可能看不到任何异乎寻常之处。”[注]普利斯特利成就颇丰，但他首先是一位化学家，而且牛顿在化学领域取得的成就似乎还佐证了这一无礼的评价。

唐突地翻开牛顿的笔记，的确会发现一些出人意料的段落。

卡德摩斯（Cadmus，腓尼基王子）杀掉的恶龙是我们研究的主题，而且他的牙齿是经过净化了的物质。

德谟克利特（Democritus，古希腊哲学家，希腊的一位智者能人）说将某些鸟类（易挥发的物质）的血与某种毒蛇（象征水星）的血混合在一起，人类饮下（并消化掉）之后就会听懂鸟类的叫声（挥发物的性质是如何固定下来的）。

信徒圣约翰（St John the Apostle）和荷马（Homer）都是智者能人。

萨克拉·贝奇［Sacra Bacchi，或称狄奥尼西卡（Dionysiaca）］，酒神巴克斯［Bacchus，或称狄俄尼索斯（Dionysus）］的仪式，由俄耳甫斯（Orpheus）创立，具有化学（Chymicall）意义。注

无独有偶，牛顿在研究化学及炼金术时写了成百上千页纸，有些是描述实验的，另一些则试图理解这种预言的深层次含义。许多人觉得牛顿最伟大的物理学作品《自然哲学的数学原理》（*Principia*）中某些段落同样晦涩难懂，这是因为这些段落都是不同推理方式的产物。最关键的差异不在于牛顿碰巧在某一案例中正确而在另一案例中错误，而在于其论据质量的变化幅度很大。这会促使极端的相对论者认为，《自然哲学的数学原理》中严格的数学论据也合理不到哪里去，还不如“海神尼普顿（Neptune）用他的三叉戟将哲学家引入学术殿堂。所以尼普顿是一种矿物质、一种水溶剂，而三叉戟则是一种类似墨丘利节杖（Caduceus of Mercury，医学的标志）的水酵素，用这种水酵素来让水银（Mercury）发酵，也就是说，用干黑铜制成的两只干白鸽”。注然而，不可否认，两种物质都是聪明大脑的产物。

牛顿推理天文学与炼金术话题时，显著的不同之处在于其获取数据的质量差异。对于天文学话题，牛顿能获取第谷·布拉赫（Tycho

Brahe）对恒星及行星位置的精确记录。而对于炼金术话题，牛顿就只能获取各类诠释性文章，只听到含糊的传闻，说人们能转换金属的属性，还有各类滥竽充数的诀窍。此外，二者还有一个不同之处。

在推理天文学话题时，牛顿知道自己必须要说服同时代的最杰出人才，还会试图预测这些人的许多反对论据。即便其学界同人不与牛顿交谈，也同样在影响牛顿思考的方式。这种关注在其出版物中有所体现。牛顿在发表其革命性观念时，就确信能说服同事，即便不能让更大范围的受众认同。虽然牛顿写《自然哲学的数学原理》的初版时“用的是大众认同的方法，以便许多人都能理解”，紧接着他意识到“若不更深层次地进行分析，便不太容易识别后果的严重性，也不能搁置人们已适应多年的偏见”。因此，“为防止这种叙述可能会遇到的争论产生”，他“选择把这本书的主旨用命题的形式（用数学的方式）加以表现”。[注]

相比之下，牛顿在研究炼金术时，缺乏严肃认真的对话者：那时，除了亡故已久的J.B.赫尔蒙特（J. B. van Helmont）、罗伯特·波义耳和他本人之外，只有“极少数‘化学哲学家’”。[注]此外，整个话题处于保密状态。就此主题，牛顿“不打算发表任何东西”，[注]他不满波义耳持有独家秘方。[注]在波义耳去世时，牛顿就问约翰·洛克（John Locke）他的遗嘱执行人有哪些，波义耳先前藏起来的笔记谁可以看，尽管与此同时，牛顿坚决否认这些笔记的价值。[注]在这种社会语境下，牛顿不必制造强有力的论据，他会反对论据的可能性也极小。

在推理地心引力时，牛顿必须要说服一群消息灵通且持怀疑态度的同行，所以不得不拿出更好的论据。在推理炼金术时，却没有这样的语境。同样才华横溢，同样是独自推理，在这一领域却一事无成。

---

1. Heisenberg 1952, p.30.
2. 引自Shapin 1991, p.194。

3. Planck 1968,pp.33-34，引自Kuhn 1962。

4. Mahoney & DeMonbreum 1977.

5. Mahoney 1977.

6. Dunbar 1995.

7. Kuhn 1962, p.152.

8. Cohen 1985, p.468.

9. 雷伊（Wray，2011）评论了这些研究，他推断出，“普朗克（Planck）的原理是荒诞的说法。年长的科学家并非特别抵制变化”（p.190）。

10. Cohen 1985, p.468.

11. 引自Cohen 1985, p.472。

12. Rorty 1991.

13. Hull, Tessner & Diamond 1978; Kitcher 1993; Levin, Stephan & Walker 1995.

14. Oreskes 1988;其他例子可参见：Kitcher 1993; Wootton 2006。

15. 引文出自Thomas 1971。

16. King 1991.

17. 引自Steven & Schaffer 1985。

18. 引自Steven & Schaffer 1985，p.75。

19. Dunbar 1995, p.380.

20. Csikszentmihalyi & Sawyer 1995, p.347.

21. Kahneman 2011, p.40.

22. Beller 2001, pp.103-104.

23. Beller 2001，pp.65-102。

24. 也可参见：Zamora Bonilla 2006。

25. Azzouni 2007.

26. Mancosu 1999.

27. Gessen 2009.

28. Westfall 1983, p.X.

29. 可参见：Fara 2002。

30. Priestley 1786, p.346.

31. 引自Hall 1996, p.188。

32. 引自Hall 1996，p.187。

33. 这里牛顿谈论的是《自然哲学的数学原理》一书的第三册。引文出自该书第三册的引言部分。

34. Hall 1996, p.196.

35. 至少这似乎是他向同事坦言的；引自Hall 1996, p.184。

36. Hall 1996, p.199.

37. Principe 2004.

结论
终为理性唱赞歌

哲学家把理性描述为一种优秀的人类心智力量。实验心理学家表明，这种超能力有重大缺陷。从进化视角来看，有缺陷的超能力形成的观念没有多大意义。所以，我们打算重新思考理性是什么，有何功能。

理性是什么（以及不是什么）

我们认为，理性是许多模块中的一个推论模块。推论模块是专门的：它们各自都有一个狭窄的职能范围，采用适合这一狭窄职能范围的程序。这一观念与老旧且仍占主导的传统观念形成对照，传统观念认为所有推论都是运用同一种逻辑（或同一种概率微积分，或者逻辑和概率的组合）。

但是，理性具有普遍性，难道这一点不是理性的特点吗？若真如此，理性怎能成为一个专门的推论模块呢？不要着急，我们答案的第一部分是要坚持认为理性确实是专门化的，仅对理由做出直觉推论。

理性是对理由进行推论吗？这可能看起来像含糊的自明之理或低廉的文字游戏（至少在英语和罗曼语中，这个源自拉丁语的词既指作为官能的理性，又指作为动因的理由）。然而，在哲学和心理学史上，理性和理由已被作为两个截然不同的话题研究过。所以，假定理性的功能是对理由进行推论，绝非自明之理，这一假定便是对主流观念的严峻挑战。可是，它是如何发挥作用的呢？

确实，无论怎样，人类都能就任何话题进行推理。只对理由进行直觉推论的机制要怎样成为理性本身的机制呢？答案的第二部分是，理性模块不仅得出关于理由的直觉结论，而且这些结论的确仅仅与理由有关。这样的话，理性模块也间接产生了对理由本身相关事物的反思结论。因为理由可以是关于任何事物的，如兔子、雨、船、人、法律或数字等，理性便可以间接制造关于任何话题的反思结论。

打个比方，若天空中有乌云，而你又想要出门，这时，你可能会凭直觉知道“可能会下雨。这是带上雨伞的强有力理由”。基于这一直觉，你决定带上雨伞，你的这一决定便是你直觉推论的间接性、反思性结论。你的直觉就是对理由说服力的直觉，你的反思性决定就是带上雨伞。

我们没有假定直觉和推理必须通过两个完全不同类型的机制来产生，这只是当前在双重过程理论中有所改善的老旧观念，相反的是，我们展示了直觉推论机制实现推理的过程。这一机制是高度专门化的，但还是对我们在各个领域的想法产生了间接有利的影响：该机制具有虚拟领域普适性。

试比较：直观知觉机制是高度专门化的。这一机制处理视网膜上由光子造成的刺激样式，并对外界可能放射或反射这些光子的事物做出无意识推论。尽管该机制高度专门化，但多亏了它，视力才对我们大多数想法及决定产生了间接有利的影响。虽然视力与理性专门化的形式迥然不同，但它们都是虚拟领域普适性的例证。

视觉程序利用了物体在正常环境中反射光线的规律性，诸如“光线一般从上面射下来”这样的事实，另外，关于各种物体类型有更多具体的规律，诸如“人脸一般看起来都是朝上的”这样的事实。同样，理性程序也会利用一般理由的属性特征，如关联性、清晰度或说服力；当然，也利用特殊类型理由的属性，如先例对协调与理由相关的问题的作用，从亲子关系到法律事宜等。

换句话说，我们展示了理性是如何与其他直觉推论模块相适应的，而不是作为束之高阁的超能力。尽管理性存在虚拟领域普适性，却不具有广泛适用性和对各种动物都有利的适应特性。我们认为，理由是为了社会消费。理性是人类对为自身建造的超社会生态位的适应特性。到此，理性之谜的第一部分便解决了。

理性有什么功能（以及没有什么功能）

我们一起研究理性已经有十多年了。虽然我们在本书中是第一次解释理性机制，但是在许多出版物及研讨会上我们已经展示过对理性功能的早期研究，即“论证推理理论”。

我们同许多哲学家及心理学家进行了交流，其中大多数人认可某一种版本的唯理智观念，他们把理性视作改善个人认知、让人独立形成更理智观念及决定的一种手段。他们认为理性理应客观且严苛。但是，当我们摆出证据，证明理性简直在无可救药地表现偏见及懒惰时，他们竟十分坦然地接受了这一观念。的确，他们许多人都熟悉该证据，有些人甚至还为收集论据做出了贡献。

读本书时，你多少会有相同的感受。当你看到贝迪永的体系或感受到鲍林对维生素C的痴迷时，可能会更倾向于找到其趣味性，而不是感到诧异。成堆的心理学书籍都在探讨人类的偏见——如同我们在第十一章至第十四章中所做的那样。这已经成为一种流行的科学文化。从某种程度上来说，我们十分清楚理性的偏见及局限——或者，至少认可他人的理性存在偏见及局限。

我们擅于发现他人持有的偏见，但不擅于承认自己的偏见。[注]或许这就解释了为何许多人既能坚持唯理智立场（既为他们自己，也为志趣相投者），又能坚持认为理性存有偏见与惰性（对不同意这些观点的人

来说尤其如此）。

实际上，唯理智理论中常见的辩护就较好地示范了存在偏见及惰性的推理。无可争辩的事实是，即便唯理智理论可取，个体推理也几乎很少（如果曾经有的话）表现出其应具备的客观与公正。我们在讨论如何处理理论与实证之间的不协调问题时，几乎不曾考虑到方法本身有误的可能性。各种影响因素及理性本身的缺点对推理的失败做出了懒惰的解释。这还是没有太大的进化意义。真正的适应特性是适应性的，真正的功能是有作用的。

从互动观念来看，我们认为理性的偏见及懒惰都不是缺陷，它们是帮助理性实现其功能的手段。人们在找寻支撑自己观点的理由时是有偏见的，因为只有这样，他们才能为自己的行为辩护，才能说服他人相信其观念。你要为自己辩护的话，就不能拿出会削弱自己立场的理由。你要说服他人改变观点的话，就不能拿出会支撑其观点的理由，也不能拿出让其摒弃你所持观点的理由。而且人们之所以懒惰地推理，是因为在典型的互动语境下，这是最有效的方式。比起费尽心力来预测反对论据，等待对话者来提供这些反对论据通常更加有效（如果他们能做到的话）。

将理性合理地理解为一种社会互动工具，这种观念当然不是最理想的，但也没什么瑕疵。到此，理性之谜的第二部分便解决了。

虽然人们已经基本认可了论证推理理论，但是还是容易对此进行错误解读，不仅批驳该理论的人如此，更令人担忧的是，被该理论吸引的人也是如此。

我们反复遇到的第一种误解是，认为论证只是操控及欺骗他人的手段，没有真正的学术价值。这种极具讽刺性的推理及论证观念一定有其吸引力——可能会让人觉得自己比幼稚的普通人厉害。可能要让一些读者失望的是，我们并不赞成这一观念，也没有这种愤世嫉俗的心态。

当然，人们有时会受到某一论据的蒙蔽。但发生这种情况仅仅是因为，与那些只是自诩权威的声明相比，大多数论据都不具有欺骗性且更容易被评估。既然不会对论据进行客观的评估，大家又为何要认真看待论据呢？推理不仅是产生说服他人的论据的工具；同样重要的是，推理还是评估他人产生的、欲说服我们的论据的工具。制造论据的能力只有与评估论据的能力并驾齐驱才能发展进化。

一方面，唯理智论者在拿出论据及评估论据时都应坚守理性的高要求及客观性。他们只能乖乖相信，人类理性实际上并不符合其理应表现的样子。

另一方面，互动理论做出了两个对比鲜明的预言。我们在提出论据时会持有偏见与惰性；在评估论据时，则应该严苛且客观——严苛是为了不受到贫乏或荒谬论据的蒙骗以致接受了错误的观念，客观则是为了在理由充分时，准备修正自身观点。

第一个预测——人们在拿出理由时存在惰性和偏见——其实不是真正的预测。我们“预测”的数据，更确切地说，倒推的数据，已经展现在众目睽睽之下。互动理论所做的（以及唯理智理论没能做到的）就是理解这一证据。

第二个预测——理由的评估是严苛且客观的——是真正的预测。我们在有关这一问题的文献中几乎找不到任何直接的证据，而且几乎没有证据是令人信服的。第二个预测源自互动理论。若让唯理智推理心理学家预测人们会更擅于提出论据还是评估论据，你的对话者很可能预测不到任何不同之处，甚至看不到该问题的依据。

有人认为，人们（至少他人）存有偏见与惰性，这一观念比较普遍。同样，认为人们轻信且固执的观念也比较普遍。轻信表现在：他们认可最明显的荒谬论据。固执表现在：他们不给合理的论据任何机会。如果人们既轻信又固执，那么传播错误的新观念就太容易了，消除老旧

的错误观念也会非常困难。如果观念交流有什么作用，那也是倾向于支持不好的观念。这种悲观的看法广为流传。特别是小组讨论经常被批判，[注]这种态度清晰地展现在一句俗语中：“委员开会，造马为驼。”

实际上，骆驼是大自然的奇迹，小组讨论通常效果非常好，而且有关这些讨论的研究间接证明了我们第二种预测的合理性。在解决问题时，小组整体表现通常要比小组中普通个人的表现好得多，而且在某些情况下，甚至比任何成员的个人表现都要好：人们能一起找出任何人都不能凭一己之力找到的解决方案。我们在第十五章中回顾了其中某些证据。[注]我们可以从文献中找到更深层次的证据来支撑我们的预言，即与提出论据相比，人们评估论据时要求更高、更客观。这一证据是间接性的。不过，我们已经开始直接检验自身预测了。拭目以待吧！

在本书中，我们凸显了小组推理的显著成效。不过，这也经常导致人们对我们的理论产生第二种误解。

根据互动理论，理性不是进化来提高独立思考能力的，而是作为一种社会互动工具。我们提出理由来为自己辩护并说服他人。这种理由交流可能对每个对话者都有好处。在某些情况下，也会对小组有好处。这样的话，为何不设想理由交流及理性机制本身的进化可能是为了小组利益，而不是为了每个个体的利益呢？

有人认为达尔文的自然选择理论在好几个层面都可以发挥作用，特别是在小组层面，该观念最近已经得到了大力发展并被进行了讨论。人们尤其认为，小组层面的选择对促使人类进行合作及形成道德观念有重要的进化意义。[注]那么，理性的进化不是应该成为小组层面认知合作选择的例子，而不是个人层面认知合作选择的例子吗？不，我们假设的肯定不是小组层面的选择。实际上，把它当作小组层面的适应特性会违背互动推理理论。

小组层面的选择更加追求集体利益而非个人利益。相比之下，理性

正如我们描述的那样，是追求个人利益的机制。个体坚持的获益方式是让他人接受其正当理由，或者拿出影响他人的论据。此外，个体还客观评估他人给出的正当理由及论据，并基于这一评估对他人的理由予以采纳或拒绝，并从中受益。在社会互动中，这些益处得以实现，但是仍然都只是个体层面的益处。

理由对互动也有作用，互动者的利益可能相同，也可能不同。理由交流可能在任何一种情况下都发挥着重要作用。例如，论证在谈判中发挥了重要作用，其中会存在利益分歧，通常还十分明显。基于小组成员利益共享的程度，他们彼此信任，而彼此信任的人之间使用正当理由及论据的频率大大减少，甚至根本用不到。在小组中，群体选择会促进该小组的系统信任及可信度。正如我们描述的那样，理性是对社会生活的适应特性，其中信任要靠自己争取，而且这种信任仍然有限且不稳定。

小组讨论并不总是有效。若人们的观念一开始就高度一致，就会导致极化现象。若开始人们的观念有冲突且目标不同，这些分歧往往会加剧。若参与者观念不同但目标相同，小组讨论通常有好处。我们认为，在这些情况下，应该把所获得的集体利益当作服务于个体利益机制的“副产品”。

我们认为理性是个体层面而非小组层面的适应特性，并不是说它只会对个体产生影响，对社群和社交网络就不会有影响。许多推理者的个人性情会对道德、政治或科学目标形成约束，我们在第十七章和第十八章中探讨了形成约束的各种方式。更普遍地说，个人性情的人口聚集效应很值得研究。在不同文化观念及惯例互动成功或失败的例子中，理性发挥着怎样的作用呢？另外，虽然我们已经展示了理性是人类的普遍共性，但是要弄清楚在各种文化传统（逻辑文化史只是该问题中一个十分有趣的层面）中如何利用、强化理性并对之进行不同的编码，任重道远。

现在，我们推断出的结论是：太长时间以来，理性站在一个满是裂

痕的基座上，悬于其他官能之上，却尴尬地倾斜着。我们希望本书能将其归位，虽然与其他认知机制平起平坐，却能使之挺直，而且要像其他进化的机制那样，用复杂且微妙的方式呈现力量，并展现无穷的魅力。

---

1. Pronin 2007.
2. Mercier, Troucheet al. 2015.
3. 例如，有数据显示，小组讨论让正确的解决方案可以完美或近乎完美地用于解决易处理的逻辑或数学问题（Claidière, Trouche & Mercier这样主张；Laughlin 2011；Trouche, Sander & Mercier 2014）。对于一些归纳问题，小组讨论让小组成员可以形成更好的集体性回答，该回答优于最好的小组成员单独想出的回答（Laughlin 2011）。小组讨论还让大中小学生在完成学校布置的各类任务时可以做得更好（Johnson & Johnson 2009; Slavin 1995）。它让陪审员可以做出更好的裁决（Hastie, Penrod & Pennington 1983）。它让科学家可以摒弃错误的假设并形成新的假设（Dunbar 1997）。它让公民可以给出更明智的建议（Fishkin 2009）。它让医生和医科学生可以做出更好的诊断（Kesson et al. 2012; Reimer, Russell & Roland 2016）。它让预报员可以做出更好的政治性及经济性预测（Mellers et al. 2014）。它让小组成员可以做出更加理性的战略决策（Kugler, Kausel & Kocher 2012）。它让投资者可以做出更好的投资（Cheung & Palan 2012）。而且，它还让好点子可以在科学家与普通大众之间传播（Chanel et al. 2011; Kitcher 1993; Wootton 2015）。
4. 例如，Boyd et al. 2003; Sober & Wilson 1998。欲了解另一种观点，可参见：Baumard, André & Sperber 2013。

致谢

自2005年起，我们几位作者针对理性这个话题开始了合作，本书正是基于此诞生的。而致力于本书的研究则开始于2010年。在这一过程中，有许多人及组织给予我们鼓励与帮助。我们以个人名义和集体名义获得了许多机会，能够在世界各地的认知科学及哲学的会议与研讨中呈现、发展和修正自己的思想理论。特别是在2011年，我们受邀前去参加伦敦大学哲学协会的Chandaria演讲，此次演讲是本书主要思想的一个较早版本。雨果曾在宾夕法尼亚大学、Cogmaster在巴黎施行的一个跨学科的学术项目和纳沙泰尔大学教授有关理性与论证的课程。2013年，基于本书做的工作，丹在布达佩斯的中欧大学的认知科学及哲学协会中教授了一门课程。许多在此难以一一列举的朋友、同事及学生，在不同场合与谈话中都给出了令我们获益匪浅的评论、建议及批评。对此，我们十分感激。

这些年来，不少同事和朋友都以个人名义给予我们很多帮助，特别是吉安-巴普蒂斯特·安德烈(Jean-Baptiste André)、尼古拉斯·宝玛德(Nicolas Baumard)、斯坦福纳·伯纳德(Stéphane Bernard)、克里斯提娜·比基耶里(Cristina Bicchieri)、莫里斯·布洛奇(Maurice Blouch)、帕斯卡尔·波伊尔(Pascal Boyer)、苏珊·凯里(Susan Carey)、科拉里·舍瓦(Coralie Chevallier)、尼古拉斯·克莱蒂亚(Nicolas Claidiére)、法布里斯·克莱门特(Fabrice Clément)、杰格·西布拉(Gergo Csibra)、丹·丹尼特(Dan Dennett)、欧菲利亚·戴罗伊(Ophelia Deroy)、纳戴因·福莱斯科(Nadine Fresco)、桂瑞·杰格力(Gyuri Gergely)、亚伯·杰士菲尔德(Abel Gerschenfeld)、克里斯托弗·海因兹(Christophe Heintz)、拉里·赫什菲尔德(Larry Hirschfeld)、皮埃尔·杰各布(Pierre Jacob)、伊莲娜·兰德默

(Hélène Landemore)、克里斯汀·朗格卢瓦(Christine Langlois)、汉娜·马诺(Hanna Marno)、奥利维尔·马斯卡罗(Olivier Mascaro)、奥利威尔·莫林(Olivier Morin)、蒂凡尼·莫里西尤(Tiffany Morisseau)、伊拉·诺威克(Ira Noveck)、格罗利亚·奥里吉(Gloria Origgi)、盖伊·波利策(Guy Politzer)、杰罗姆·普拉多(Jérôme Prado)、路易斯·德·萨雪(Louis de Saussure)、汤姆·斯科特-菲利普斯(Thom Scott-Phillips)、巴里·史密斯(Barry Smith)、布莱恩特·斯特克兰德(Brent Strickland)、丹尼斯·塔托恩(Denis Tatone)、拉多·昂布拉斯(Radu Umbres)、吉安-巴普蒂斯特·凡·德尔·亨斯特(Jean-Baptiste Van der Henst)、戴尔杰·威尔逊(Deirdre Wilson)和赫罗什·亚玛(Hiroshi Yama)。感谢你们！

维托里奥·吉罗托（Vittorio Girotto）是当代理性研究的领军人物，他以朋友的身份给予我们鼓励的同时，也同我们并肩作战，令人大受鼓舞。他于2016年4月逝世，享年59岁。我们对他的逝去抱以万分的想念。

对于看过本书草稿或部分草稿并给出建设性意见的同事，我们也非常感激：克拉克·巴瑞特（Clark Barrett）、汉诺克·本-亚米（Hannoch Ben-Yami）、帕斯科·博伊尔（Pascal Boyer）、彼得·卡鲁瑟斯（Peter Carruthers）、乔治·加法尔维（George Gafalvi）、皮埃尔·杰各布、菲利普·约翰逊-莱尔德、奥利威尔·莫林、伊拉·诺威克、约瑟夫·珀纳、菲利普·佩蒂特（Philip Pettit）、汤姆·斯科特-菲利普斯、约翰·沃森（John Watson）、戴尔杰·威尔逊、赫罗什·亚玛，以及哈佛大学出版社一名匿名评论家。

我们对代理人约翰（John）及麦克斯·布罗克曼（Max Brockman）、编辑伊恩·马尔科姆（Ian Malcolm）及哈佛大学出版社团队表示由衷的感谢。

雨果受益于以下机构或项目：法国武器装备总署（Direction Générale de l'Armement）的资金支持［特别是迪蒂亚·巴扎尔盖特

（Didier Bzazalgette）］；宾夕法尼亚大学的哲学、政治学及经济学项目；斯蒂文·F.戈德斯通（Steven F. Goldstone）慷慨的支持；瑞士纳沙泰尔大学的认知科学研究团队、瑞士国家科学基金会（Swiss National Science Foundation）；最后，是他现在工作的地方——法国国家科研中心（Centre National de la Recherche Scientifique），他已加入马克·詹尼诺德认知科学协会（Institut des Sciences Cognitive Marc Jeannerod）。雨果对本书实证研究部分所做的大部分贡献，都是与其两位才华横溢的博士生协作完成的，他们是托马斯·卡斯特兰与伊曼纽尔·特鲁奇（Emmanuel Trouche）。他最感谢的是自己的两个家庭——一个是自他还是个年轻学生之时就开始支持着他，陪他走过了多年学术漫游生涯（并且现在仍一如既往），以及现在支持着他的家庭——特雷泽（Therese）、克里斯托弗（Christopher）与亚瑟（Arthur）。事实上，雨果应当把更多的工作时间用以陪伴家人。

丹对以下机构的成员、职员及学生表示感谢：布达佩斯中欧大学的认知科学及哲学系和巴黎吉安·尼柯德协会（The Institut Jean Nicod）。其研究还得到欧洲研究协会（European Reaserch Council）的拨款协助［款项编号：ERC-2013-SyG，建构社会思想：交流，协作与文化传播，ERC，款项合同编号（609619）］。这些年来，儿子南森（Nathan）与利奥（Leo）给予了他莫大的快乐。